ANÁLISIS CUALITATIVO COMPARATIVO
TOMO I

NÍTIDO (csQCA)
Teoría y Práctica

Juan Mejía Trejo



AMIDI
Academia Mexicana
de Investigación y Docencia
en Innovación

ANÁLISIS CUALITATIVO COMPARATIVO

TOMO I

NÍTIDO (csQCA)

Teoría y Práctica

Juan Mejía Trejo



Este libro fue sometido a un proceso de dictamen por pares de acuerdo con las normas establecidas por el Comité Editorial de la Academia Mexicana de Investigación y Docencia en Innovación (AMIDI)

Primera edición, 2023

Av. Paseo de los Virreyes 920.
Col. Virreyes Residencial
C.P. 45110, Zapopan, Jalisco, México

ISBN Tomo I: 978-607-59567-1-8
ISBN Obra Completa: 978-607-59567-0-1

Índice

Introducción

El objetivo de esta obra es la de introducir al lector en los conceptos básicos del uso del Análisis Cualitativo Comparativo (**QCA.** *Qualitative Comparative Analysis)* con el conjunto de *datos nítidos* (*crisp-sets.***csQCA**) a través de demostraciones de su uso con ejemplos varios en la administración de la innovación.

Desde fines de los años 80, su precursor, el análisis cualitativo comparativo **QCA** (*Qualitative Comparative Analysis*) ha estado en el centro de atención de la metodología de las ciencias sociales ya que se fundamenta, en el conjunto de relaciones y objetivos establecidos del descubrimiento de condiciones *suficientes* y *necesarias*.

Cabe señalar, que es común referirse a la versión Booleana original del **QCA** como **csQCA**, donde **cs** (*crisp-sets*) denota un conjunto de *datos nítidos*, la versión que permite utilizar las condiciones de múltiples categorías, se denomina como **mvQCA**, donde **mv** (*multi-value*) describe el valor múltiple y para la versión **fsQCA** donde **fs** (*fuzzy-set*) describe al conjunto de *datos difuso*. El objetivo del análisis de **QCA**, en general, es dar cuenta de un resultado particular, contra los análisis basados en *regresión*, que por el contrario, generalmente tienen como objetivo ser herramientas base para *explicar los efectos de las causas* **(Wagemann y Schneider, 2010).**

Sobre el estado de las ciencias sociales, como base general de la administración de la innovación, **Sartori (1970),** llegó a afirmar:

> *"...se aprecia un pésimo estado de la ciencia oscilando entre dos extremos poco sólidos: el pensamiento inconsciente, lo que hace una abrumadora mayoría, y un pensamiento demasiado consciente, hecho por una pequeña minoría..."*

Llamando a los estudiosos a adquirir capacitación en lógica (*primaria*):

> *"Para dirigir un curso intermedio entre mal manejo lógico crudo por un lado, y perfeccionismo lógico (y parálisis) por otro lado…"*

Fue a fines de la década de 1980, cuando el Dr. Charles Ragin trajo el álgebra Booleana y la teoría de conjuntos, para las ciencias sociales con su innovador libro *The Comparative Method* **(Ragin, 1987)** que describe a profundidad todo lo relativo a los componentes del **QCA**. Aún así, el verdadero estímulo en la atención comenzó algunos años después, con el libro *Fuzzy-sets Social Science* **(Ragin, 2000).** Por ahora, ya hay académicos que usan los *métodos comparativos configuracionales,* dada la posibilidad en formalizar el *análisis orientado a casos* y de este modo, ofrecer herramientas para mejorar la investigación comparativa. Estos métodos son particularmente aptos para identificar lo *mínimamente necesario* y/o *mínimamente suficiente* (*combinaciones de*) condiciones que producen un *resultado de interés* (es decir, *evaluar las causas de los efectos*), con gran potencial de aplicaciones cuantitativas como en la ingeniería **(Mendel y Korjani, 2010; Marks, et al., 2018).**

Dadas las ventajas que se ofrecen, en discusiones recientes sobre los métodos comparativos *configuracionales*, los estudiosos sostienen que los enfoques cuantitativos basados en regresiones *vs.* el **QCA**, se aplican mejor junto a otro **(Ragin 2008; Schneider y Wagemann 2010; Rihoux 2006).** Sin embargo, existe una advertencia para los entusiastas, de que los académicos no deban convertirse en *monomaníacos* del **QCA** (**Ragin y Rihoux 2004, p. 6**).

Por otro lado, se tienen primeros esfuerzos de la aplicación de **fsQCA,** en el área del emprendimiento y la innovación como el trabajo de **Kraus et al., (2017)** donde se hace una recopliación de **77** artículos publicados de 2005 a 2016 con las palabras clave de: **fsQCA**, administración de negocios, empredimiento e innovación, revelando un incremento paulatino en estos campos para el uso de la **fsQCA**. Es así, que las ciencias de la administración orientadas a la innovación, tienen la posibilidad de aprovechar lo

realizado en las ciencias sociales a través de las importantes aportaciones del Dr. Ragin, por lo que la presente obra, se compone de once capítulos, los cuales describimos brevemente:

CAPÍTULO 1. Análisis de regresión vs. análisis cualitativo comparativo de datos (QCA). El capítulo presenta cómo funciona la regresión y su uso por parte de los investigadores sociales orientados a la administración de la innovación; los errores más usuales en los que se incurre y por lo tanto, las inconveniencias para la investigación. Esto trae como consecuencia potenciales errores al asumir hipótesis que al ser sometidas al **QCA**, resta mucho su efectividad. El capítulo cierra con la descripción de la **fsQCA,** que permite visualizar su incursión en las ciencias de la administración de la innovación.

CAPÍTULO 2. Software QCA y configuración fsQCA. Este capítulo, permite al lector conocer los alcances del software **fsQCA**, diseñado por el Dr. Charles Ragin a fin de conocer cada una de las funciones que lo componen desde crear y abrir un archivo en diferentes formatos, hasta configurarlos y probar las funciones disponibles, en doce ejercicios prácticos con el uso del software **fsQCA** que se describen a detalle.

CAPÍTULO 3. Análisis Cualitativo Comparativo de Datos Nítidos (csQCA). Este capítulo reporta al lector un tipo de datos muy común en el uso de la **csQCA**, revelando conceptos básicos de su uso a través de el uso de datos binarios, la negación Booleana, el uso de tablas de verdad, los agrupamientos, la adición y multiplicación Booleanas, la lógica combinatoria, las implicaciones principales, la Ley de De Morgan, las causas necesarias y suficientes y la minimización. Trata los conceptos acerca de los casos contrafactuales, el problema de los casos coincidentes, cómo se comparan los casos contrafactuales vs. **QCA** y un contraste con la investigación convencional cuantitativa. Se aplican seis ejercicios con el uso del software **fsQCA** que confirman el uso de los principales conceptos.

CAPÍTULO 4. Análisis Cualitativo Comparativo de Datos Difusos (fsQCA). Este capítulo reporta al lector un tipo de datos muy especial, el de tipo difuso que usa el software **fsQCA**, revelando conceptos básicos de su uso a través del conocimiento de su naturaleza, la posibilidad de usarlos

por niveles y de manera contínua. Se presentan operaciones del conjunto de *datos difusos*, tales como: la negación Booleana, la conjunción (lógica **AND**), la unión (lógica **OR**). Incluye modelos y métodos de calibración análisis de condiciones necesarias y suficientes así como de consistencia, cómo realizar tablas de verdad y analizar las esquinas de espacio vectorial. Se aplican catorce ejercicios que demuestran el uso del software **fsQCA** que confirman el uso de los principales conceptos.

CAPÍTULO 5. Evaluación del conjunto de datos fsQCA. El capítulo hace una descripción de la importancia de lo que se conoce como condicones **INUS** y **SUIN** así como de *consistencia y cobertura* en general. Se presenta el desarrollo manual de cinco casos que confirman el uso de los principales conceptos.

CAPÍTULO 6. La calibración y su importancia en fsQCA. Este capítulo abre un debate sobre las implicaciones de la calibración antes de iniciar mediciones en alcance y contexto, los pros y contras de utilizar indicadores por parte de la investigación cuantitativa, relacion de **SEM** vs. **fsQCA**, la investigación cualitativa y la necesidad de la calibración, los métodos directo e indirecto de calibración. Se presenta el desarrollo manual de cuatro casos que confirman el uso de los principales conceptos.

CAPÍTULO 7. Pensamiento configuracional en fsQCA. Este capítulo demuestra, la importancia del pensamiento *configuracional* desde el punto de vista cualitativo y cuantitativo, se discute sobre la evaluación del grado de pertenencia en una configuración así como la comparación de trayectorias causales.

CAPÍTULO 8. Efectos netos en fsQCA. El capítulo describe lo que se debe entender por *efectos netos*, los problemas que se tiene del mismo, se propone un cambio de enfoque a tipos de casos, se hace exposición de comparación de casos de *datos difusos* y el análisis de *configuraciones*.

CAPÍTULO 9. Efectos netos vs. Configuraciones en fsQCA. A través de dos casos hipotéticos, se realiza el estudio y se desarrollan cada uno de los conceptos vistos explicando por comparación sus resultados, a fin de comprender la naturaleza de los efectos netos de los de configuración por **fsQCA**.

CAPÍTULO 10. Guía de análisis csQCA. Con datos *nítidos*, de un caso de innovación, se ofrece al lector una guía rápida y visual, a través de *screen-shots* y 8 ejercicios de cómo resolver un caso a través del análisis de condiciones necesarias, representar datos en la tabla de verdad, identificar contradictorios y remanentes lógicos, análisis de suficiencia, minimización de la tabla de verdad para ocurrencia y no ocurrencia del fenómeno e interpretación de resultados.

CAPÍTULO 11. Guía de análisis fsQCA. Con datos *difusos*, de un caso de innovación, se ofrece al lector una guía rápida y visual, a través de *screen-shots* y 9 ejercicios, cómo resolver un caso a través del acceso al archivo de trabajo, el análisis de condiciones necesarias con ocurrencia del fenómeno y sus reportes, la interpretación de la cobertura, la generación de la tabla de verdad, la aplicación de la opción *Standard Analyses*, interpretación de las soluciones generadas, la generación de la tabla de verdad con no ocurrencia del fenómeno, el análisis de implicaciones principales e interpretación del caso.

GLOSARIO. Se presenta un glosario de los términos más utilizados al respecto.

CAPÍTULO 1.

Análisis de regresión vs. análisis cualitativo comparativo de datos (QCA)

En el estudio de las ciencias sociales, al aplicar cualquiera de los enfoques de la investigación conocidos cualitatitivo o cuantitativo, tendemos a realizar una previsualización de los resultados a obtener a partir de diseñar el tipo de datos que se van a recopilar. Por ejemplo, si es la investigación de enfoque cuantitativo, los datos de tipo nominal es factible hacerlos escalares y practicarles técnicas de predicción basadas en regresiones o si la investigación es de enfoque cualitativo, es posible a partir de los datos de tipo nominal, realizar con ellos categorías del fenómeno a tratar.

Una actualización en el análisis de regresión

Es importante estar consciente de que toda toma de decisiones debe hacerse con datos, y que, por lo tanto, debe de tener claro la técnica que va a utilizar para analizarlos, así como su interpretación. Uno de los más importantes tipos de técnicas de análisis multivariante con enfoque dependiente, basadas en la correlación, es el análisis de regresión.

Para los investigadores sociales con orientación a la administración de la innovación con enfoque cuantitativo, el análisis de regresión representa una técnica de clasificación matemática que revela cuál variable es la que produce mayor impacto. Responde a preguntas típicas como: ¿qué variables son las más importantes?, ¿cuáles variables son las que se deben ignorar? Y tal vez, la más importante ¿qué certidumbre se tiene acerca de dichas variables?

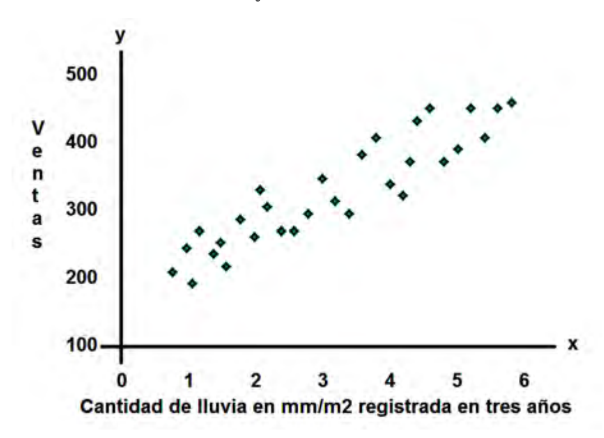
Así, se tienen una serie de variables independientes que afectan a una dependiente y aplica por ejemplo a casos como, ¿qué variables son las que afectan la venta de cierto producto o servicio mensualmente?, ¿cómo se comporta la variable de la propuesta de valor en un modelo de negocios

respecto al resto de las variables que lo componen? O ¿cómo afectan la cultura, el tamaño de la organización, la incertidumbre del entorno, el tipo de servicios prestados a la innovación tecnológica que genera una empresa de telecomunicaciones? (Ver **Mejía-Trejo, 2019a** en el Capítulo 5).

¿Cómo funciona el análisis de regresión?

Para llevar acabo un análisis de regresión, se deben capturar las variables independientes en cuestión, por ejemplo, suponga que usted está analizando un negocio del que tiene registradas las ventas, como variable dependiente, de los últimos tres años y que desea asociarlas digamos, con la variable independiente, cantidad de lluvia registrada en dicho periodo (variables independientes) y que producen, por tanto, la **Gráfica 1.1.**

Gráfica 1.1. Construyendo la relación de variables



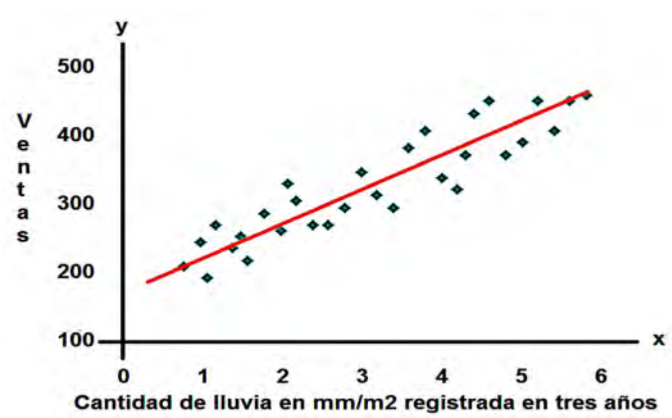
Fuente: Elaboración propia.

Como se observa, el eje Y representa a las ventas de los últimos tres años (la variable dependiente siempre se encontrará en dicho eje) y el eje X, la cantidad de lluvia registrada en dicho periodo. Cada punto en la gráfica representa el cruce de ventas con la cantidad de lluvias registrado durante los tres años en cuestión. Si observa con detenimiento los datos, probablemente se de cuenta que las ventas son más altas en cuando llueve más y

aunque es interesante saberlo, la pregunta es ¿qué tanto?, por ejemplo, si cayera una cantidad de lluvia por ejemplo de 1 mm/m^2 ¿cuántas unidades de ventas se tendrían? y ¿si aumentara a 6 mm/ m^2?

Para responder lo anterior, imagine dibujar una línea a través del gráfico que atraviese justamente por la mitad de la distancia entre los puntos graficados. Esta línea le ayudará a responder con cierto grado de certeza, qué tanto usted vende de acuerdo con la cantidad de lluvia registrada. Esta línea es llamada *línea de regresión* (es posible graficarla en varios paquetes de software con funciones estadísticas como el **Excel**, **SPSS** o **STATA**), es la línea que mejor se ajusta a los datos; es decir, *es la línea que ofrece la mejor explicación de la relación entre las variables independientes y la variable dependiente.* Ver **Gráfica 1.2.**

Gráfica 1.2. Construyendo el modelo de regresión



Fuente: Elaboración propia.

Además de producir los gráficos, el programa estadístico le arroja las fórmulas que explican la pendiente de la línea en expresiones como:

$$Y = 150 + 10X + error$$

El término de error, se refiere al hecho de que la regresión no es precisamente perfecta debido a los ajustes para crear la *línea de regresión*. Si solo se toma en cuenta la expresión:

$$Y = 150 + 10X$$

Esta expresión nos dice que si **10X = 0**, entonces **Y = 150.** Así que, refiriéndonos a los datos recolectados históricamente, cuando no se presentan lluvias, se obtienen cuando menos 1**50** unidades de venta esperando que continúe así en el supuesto de que las otras variables permanezcan igual. Así también y tomando en cuenta los datos históricos, por cada incremento de lluvia en mm/m^2 se obtienen en promedio un adicional **10** unidades de venta, esto es: por cada incremento unitario de **X** se producen **10** de **Y**. Volviendo al término de error, se tiende a decir que: la lluvia tienen un gran impacto en las ventas por cada mm/m^2 se consiguen **10** unidades de venta, pero si esta variable merece su atención, dependerá del término de error.

Una *línea de regresión* tendrá un *término de error*, debido a que, en la vida real, las variables independientes nunca son *variables predictoras* perfectas de las variables dependientes. Más que una línea rigurosa, es un estimado de los datos disponibles recolectados, así que el término de error nos informa qué tan segura puede ser su la fórmula de su recta. Mientras más grande sea, menos segura es su fórmula de la recta. La fórmula anterior, utiliza solo una variable para predecir la variable dependiente de su interés (en este caso, cantidad de lluvia que predice las ventas de paraguas).

Normalmente, los investigadores arrancan con el análisis de regresión deseando entender el impacto de varias variables independientes. En este caso, se puede incluir no solamente datos de la lluvia, sino también, los datos de publicidad de los competidores. De hecho, se puede continuar incluyendo variables independientes tanto como el término de error permanezca pequeño, de tal forma que la línea de regresión siempre se ajuste a los datos. Siempre existirán peligros al incrementar las variables independientes en un análisis de regresión en la que solo los analistas más hábiles, podrán sugerir cómo minimizar los riesgos. Es así, que podemos afirmar que la consideración del impacto de múltiples variables independientes al

mismo tiempo, sobre una variable dependiente, es una de las más grandes ventajas que tiene la técnica del análisis de regresión.

¿Cómo usan los investigadores el análisis de regresión?

Los investigadores sociales con orientación a la administración de la innovación con enfoque cuantitativo, consideran al análisis de regresión como una de las técnicas analíticas más efectivas. De hecho, es utilizada en tres vertientes:

1. Para explicar los fenómenos que se requieren entender (por ejemplo, hallar el porqué las variables de regulación legal no han producido las innovaciones tecnológicas esperadas). Se debe recordar por cierto, que la correlación no es la causalidad (*correlation is no causation*, **Mejía-Trejo, 2019a** en el Capítulo 5)**;** así que cada vez que trabaje con análisis de regresión que trate de explicar el impacto de uno o varias variables sobre otra, debe de recordar esta nota. Esto es crítico, ya que el hecho de comprobar, en el ejemplo de las **Gráficas 1.1 y 1.2,** que existe una correlación de la lluvia con las ventas, no quiere decir, que la lluvia sea la causa de las ventas, a menos de que el negocio en análisis sea de venta exclusiva de paraguas, sería muy difícil probar lo anterior.

2. Como herramienta de predicción (por ejemplo, cómo lucirá una región en el desarrollo de negocios tecnológicos a un plazo de meses o años si a partir de hoy, se impulsa la innovación abierta).

3. O decidir qué hacer (por ejemplo, incentivar las innovaciones de proceso, las innovaciones por mercadotecnia o las innovaciones del modelo de negocios para impulsar la innovación de servicios en cierto clúster).

En algunas ocasiones, variables que obviamente no están conectados por causa y efecto, sí se encuentran relacionadas (por ejemplo, en la innovación del modelo de negocios). Cuando usted observe una correlación que parta del análisis de regresión, se sugiere tener objetividad y evitar hacer supuestos. Como investigador, se recomienda revisar su marco teórico de la

investigación y más aún, la posibilidad de comprobar que está sucediendo en el mundo real, cuestionándose por ejemplo, si se trata de la innovación por mercadotecnia, cómo se manifiesta el compromiso del cliente hacia su producto (*customer engagement*) cuando llueve y no llueve, o si se trata de innovación de producto, qué tipo de mecanismo físico hace que los consumidores opten por la compra del producto cuando llueve y no llueve o si se trata de innovación de servicios, incluso, hablar con los consumidores (*focus group*) y detectar qué es lo que sienten ellos (*customer sentiment*) cuando adquieren el servicio cuando llueve y no llueve. Desafortunadamente, existe investigación académica y profesonal que se salta estos pasos por considerarlos obvios y de escasa utilidad. Lo importante no es determinar qué está sucediendo con los datos, sino, qué está sucediendo en la realidad. Use los datos para guiar más experimentos y no proponga *conclusiones aceleradas* para determinar la causa y el efecto del fenómeno en estudio **(Gallo, 2017)**.

Errores más usuales en que los investigadores caen

Dada la importancia de la técnica y si usted es un investigador con alta tendencia a su uso, se recomienda seguir, los pasos de la **Tabla 1.1**

Tabla 1.1. Pasos recomendados a seguir para el uso del análisis de regresión

Paso	Descripción
1	*Describa lo más posible su variable independiente a partir del marco teórico de la investigación*. En este caso, usted no puede cambiar cuanto llueve, pero ¿qué tan importante es esto en el fenómeno de analizar sus ventas? De hecho, no se puede hacer nada al respecto de las acciones de la competencia cuando llueve (o no llueve) pero sí es posible realizar acciones por digamos, innovación organizacional de la firma para elevar su efectividad.
2	*Evite sesgarse*. No se mentalice en descubrir, ni sesgue a sus colaboradores como investigador, qué es lo que está causando el fenómeno (en este ejemplo, no se mentalice en descubrir qué está causando las ventas). La forma enw que la mayoría de los análisis se llevan a cabo, permite entrever que se realizan sin un enfoque riguroso de lo que se está buscando. Documente su marco teórico de investigación para que sea más enfocada y precisa su búsqueda ya que su trabajo como investigador es identificar las variables de las que sospecha que son de impacto y solicitar a sus colaboradores que los midan. No les pida que ellos busquen dichas variables que pudiesen analizar ya que encontrará relaciones que muy probablemente no existen.
3	*Los análisis de regresión son extremadamente sensibles a los datos malos*. Sea muy cuidadoso con los datos a recolectar, así como la forma en que lo hace y una vez identificadas los datos atípicos, ausentes y/o con errores, una vez suprimidos, corregidos, por sustituídos o imputados, decidir si son confiables o no. Recuerde que no deberá ser más del 10% de los datos intervenidos. (**Mejía-Trejo, 2019a,** en el Capítulo 3). Así cabe preguntarse: ¿qué tan seguro se encuentra de sus correlaciones?, ¿los beneficios de la acción superan los riesgos? Vea la **Gráfica 1.3.**
4	*No ignore los términos de error.* Este es un error fatal al que tienden los investigadores con escasa o nula experiencia en el uso del análisis de regresión, ya que establecen relaciones entre variables como si fueran legítimas, certeras. Se debe recordar que los resultados aún son inciertos. Si los resultados encontrados de un análisis de regresión explican el 90% de la relación, está bien, pero si solo explica el 10% de la relación y usted actúa como si explicara el 90% de la misma, está definitivamente cayendo en error. El punto central del análisis de regresión es calcular la certeza en que el fenómeno ocurrirá. En este ejemplo, no se trata de afirmar qué cantidad de lluvia influenciarán las ventas, sino de afirmar la probabilidad de lluvia que influirá en las ventas.

5	*No deje que los datos influencíen en su intuición.* A partir del marco teórico de la investigación y su experiencia, debe ubicar su intuición por encima de los datos. Pregúntese siempre si los resultados se ajustan a lo que su conocimiento, entendimiento y experiencia tienen del contexto de la situación. Si usted detecta algo que no le haga sentido, pregúntese si los datos cumplen de acuerdo con lo indicado en el punto **3** o si existe de hecho, un gran término de error presente. Se recomienda revisar el marco teórico de la investigación nuevamente y/o consultar con colegas expertos u otros análisis de corrección de datos para estar seguro de la fuente de lo que no le hace sentido. Nunca olvide ver más allá de los números, tome en cuenta lo que está sucediendo en la realidad.

Fuente: Gallo (2017), con adaptación propia.

Gráfica 1.3. Cuándo actuar en la correlación de sus datos

Confiabilidad
Correlación frecuente, con clara hipótesis causal
Correlación frecuente, pero con muchas hipótesis causales
Correlación infrecuente e inestable
ACTUAR
NO ACTUAR
Los riesgos superan beneficios
Los beneficios superan riesgos
Beneficios de la acción en relación con el costo de estar equivocado

Fuente: Gallo (2017) con adaptación propia.

Inconveniencias para la investigación social

Desafortunadamente, la práctica del enfoque correlacional, de forma unilateral pierde lo que se llama el análisis de las relaciones en la investigación social **(Ragin, 2008)**. Lo anterior, tiene su origen en que por lo general, la teoría de las ciencias sociales con orientación a la administración de la innovación se transmite de manera verbal y que al probarse, se formula en términos establecidos *(terms of sets)* o de las relaciones establecidas (*set relations*) (*Ibídem*).

Es así, que si se expresa las pymes son reactivas al invertir en innovación, tal aseveración es traducida por los científicos sociales con orienta-

ción a la administración de la innovación como una hipótesis basada en cómo las PYMES (variable dependiente) se relacionan con la reactividad al invertir en innovación, a partir de una serie de variables independientes con la práctica de análisis multivariante con enfoque dependiente, particular-mente con técnicas de correlación (**Mejía-Trejo, 2019a**, en el Capítulo 5). Sin embargo, el practicar este tipo de técnicas deja de lado la explicación que aportan los propios términos, es decir, las declaraciones sobre las rela-ciones descubiertas, o establecidas, no de sus correlaciones, visión que es más integral en la explicación de las ciencias sociales con orientación a las ciencias de la administración.

Se dice que lo más simple y básico de una *relación establecida (set relation)* es su subconjunto *(subset)* y que son fáciles de entender e interpretar cuando se involucran categorías anidadas. Por ejemplo, los canarios, son un subconjunto de las aves; la penicilina es un subconjunto de los antibióticos; los Testigos de Jehová son un subconjunto del cristianismo y estos son un subconjunto de las religiones monoteístas o la difusión de la innovación es un conjunto de la innovación del modelo de negocios que a su vez, es un subconjunto de la innovación (OECD, 2018) etcétera. Como se observa, el hecho de que los primeros subconjuntos (canarios a aves; penicilina a antibióticos), se entienden de forma directa, se asumen por lo tanto, claros o nítidos, por lo tanto, válidos.

De hecho, son susceptibles de representar mediante diagramas de conjunto como los de Venn. Sin embargo, los casos como el de los Testigos de Jehová o la difusión de la innovación, no son claros y por lo tanto, son considerados difusos (*fuzzy*). Por otro lado, más importante que la definición de dichos subconjuntos, se encuentra la definición de las subrelaciones establecidas (*subset relations*), que describen los fenómenos sociales que están conectados causalmente o de alguna foma integral.

Así, cuando los investigadores sociales orientados a la administración de la innovación afirmen que solo las grandes empresas transnacionales creadoras de tecnología, son capaces de generar innovación se establece que, como especialistas consideran, que las grandes empresas transnacionales creadoras de tecnología forman un contundente subconjunto del total de las

empresas que son capaces de generar innovación, argumentando aún más, que las empresas transnacionales son la causa de la innovación. Igualmente, un investigador social orientado a la administración de la administración y especializado en políticas públicas que argumente que tener un fuerte cuerpo legislativo es un factor necesario o esencial de ser un país desarrollado implica que dichos países constituyen un subconjunto consistente de aquellos con un fuerte cuerpo legislativo. De esta forma, se observa que la conexión está radicada en la palabra constituyen, como opuesto a causal.

Cuando el conjunto de datos establecidas (*set relations*) reflejan conexiones sociales o causales integrales y no son claras en su naturaleza, éstas requieren explicación, por lo tanto, son dependientes de un marco teórico de la investigación basado en el conocimiento. Suponga por ejemplo, que entre las pymes que han adoptado procesos de innovación estas fallan en su competitividad. Por lo tanto, las pymes con procesos de innovación forman un subconjunto de empresas que fallan en su competitividad; así, se puede preguntar ¿fueron las fallas en la competitividad de las pymes que han adoptado procesos de innovación una coincidencia o tienen una causa sistemática? O ¿existe alguna causa o algún tipo de conexión integral entre las pymes que han adoptado procesos de innovación y su subsecuente falla? La última pregunta, hace ver un conjunto de datos establecidas (*set relations*) que sus conexiones son basadas en un marco teórico de la investigación, que requieren explicación, no son obvias. Este tipo de conjunto de datos establecidos (*set relations*), deberá explicarse basado en las teorías de las ciencias sociales especialmente en las orientadas a la administración de la innovación.

Conexiones basadas en el marco teórico de la investigación

Un importante aspecto de las conexiones basadas en el marco teórico de la investigación, en oposición a las conexiones basadas en correlaciones, es que son *asimétricas* **(Ragin, 2008)**. Esto es, por ejemplo, el hecho de que existan muchos CEOs de pymes innovadoras, quienes no tienen educación

superior, no implica de ninguna forma, que, al no tener educación superior, los CEOs sean de pymes innovadoras.

Otro ejemplo sería, que las empresas transnacionales son innovadoras, en esencia, implica que el conjunto de dichas empresas es un subconjunto del conjunto de las que son innovadoras. El hecho de que existan empresas que no son transnacionales que también son innovadoras no socava la afirmación inicial de que las empresas transnacionales son innovadoras. Una versión simétrica de esta afirmación sería: las empresas transnacionales son innovadoras y las no transnacionales no son innovadoras. Sin embargo, esta reformulación del argumento se extiende en formas que no son garantizables para su validación, esto es, el argumento es asimétrico, como lo son las formulaciones basadas en el marco teórico de la investigación.

El error de asumir hipótesis correlacionales

Los argumentos basados en el marco teórico de la investigación son erróneamente y de manera frecuente, formuladas como hipótesis correlacionales. Este error es, de hecho, uno de los más comunes en la práctica de las ciencias sociales con orientación a la administración de la innovación contemporáneas **(Ragin, 2008)**. Por ejemplo, una teoría de la innovación puede afirmar que, debido a la falta de claridad de los gobiernos, al otorgar estímulos para detonar innovación para las pymes innovadoras de reciente creación, son las que tienden a ser las menos exitosas.

Un científico de las ciencias sociales con orientación a la administración de la innovación, al leer dicho argumento, intentaría probarlo al examinar la correlación entre estímulos para detonar innovación y menos exitosas usando los datos de las pymes innovadoras de reciente creación. Suponga de nuevo, que la evidencia del marco teórico de la investigación soporta la teoría: las pymes innovadoras de reciente creación que son estimuladas por los gobiernos para detonar innovación son un subconjunto de pymes que tienden ser menos exitosas. A pesar de esta clara aseveración, la correlación entre estímulos para detonar innovación y menos exitosas usando los

datos de las pymes innovadoras de reciente creación, puede ser todavía muy débil.

Esto es debido al hecho de que hay muchas otras causas de la falta de éxito que son debidas a instancias diversas en las pymes innovadoras de reciente creación tal como la falta de un plan de negocio o un erróneo cálculo de precio de producto y/o servicio. De hecho, la afirmación que parte del marco teórico de la investigación: las pymes innovadoras de reciente creación sin un plan de negocios tienden a ser menos exitosas, no se refuta en estos casos, pero, las pymes innovadoras de reciente creación con un erróneo cálculo de precio de producto y/o servicio, tienden a ser menos exitosas pueden socavar la correlación entre estímulos para detonar innovación y menos exitosas. La diferencia clave entre conexiones basadas en correlaciones vs. las conexiones basadas en el marco teórico de la investigación, se ilustran en los casos hipotéticos de las **Tablas 1.2 y 1.3.**

Tabla 1.2. Caso hipotético: conexión basada en conexiones por correlación

Estatus	Falta plan de negocio	Cálculo erróneo de precio en producto/servicio
Pyme innovadora menos exitosa (falla)	7	11
Pyme innovadora exitosa (sobrevive)	17	5

Fuente: Elaboración propia.

Tabla 1.3. Caso hipotético: conexión basada en el marco teórico de la investigación

Estatus	Falta plan de negocio	Cálculo erróneo de precio en producto/servicio
Pyme innovadora menos exitosa (falla)	15	16
Pyme innovadora exitosa, (sobrevive)	9	0

Fuente: Elaboración propia.

La **Tabla 1.2** muestra un modelo de resultados consistentes con la existencia de una conexión correlacional entre estímulos para detonar innovación y menos exitosas entre las pymes innovadoras de reciente creación. La primera columna muestra la tendencia de las pymes innovadoras por ser menos exitosas y las mismas por ser exitosas. Si bien este resultado es muy satisfactorio desde un punto de vista de conexión correlacional, esta tabla sería insatisfactoria para un investigador interesado en conexiones teóricas de conjuntos, ya que no hay conexiones en la tabla que pudieran describirse como explícitas o consistentes. La **Tabla 1.3** sin embargo, sería de gran interés a dicho investigador debido a que muestra una conexión consistente entre digamos el cálculo erróneo de precio en producto/servicio y la *pyme innovadora* menos exitosa (los **16** casos con esta opción fallaron como se observa en la segunda columna de la **Tabla 1.3**). Mientras es significante al investigador interesado en conexiones basadas en el marco teórico de la investigación, esta tabla deberá decepcionar a aquel que sea interesado en las conexiones correlacionales, por ejemplo, entre las formas de planes de negocio, cálculo de precio y el grado de medición de las pymes exitosas; éstas todavía pudieran parecer aún débiles. Vea la **Tabla 1.4.**

Tabla 1.4. Etapas de cómo ubicar las relaciones de conjunto

Item	Etapa
1	Involucramiento de conexiones causales o de otro tipo ligadas al fenómeno social (no solo conceptual).
2	¿Es la teoría y el conocimiento dependiente?, es decir, que requiera explicación.
3	¿Está centrado a la teoría de las ciencias de la administración?, esto es, debido a que la teoría se genera básicamente de manera verbal, éstos deben de proceder de un marco teórico de la investigación.
4	Los argumentos teóricos, por lo tanto, son asimétricos, por lo tanto, no deben ser reformulados como argumentos correlacionales.
5	El resultado puede ser muy fuerte a pesar de tener correlaciones modestas.

Fuente: Ragin (2008) con adaptación propia.

Relaciones basadas en el marco teórico de la investigación

En la investigación de las ciencias sociales con enfoque a las ciencias de la administración, tenemos investigaciones que son *orientadas a caso* (así como los investigadores que se aplican al enfoque cualitativo) y están centralmente preocupados con el análisis de las relaciones establecidas (*set relations*) cuyos esfuerzos se orientan a identificar conexiones explícitas **(Ragin y Rihoux, 2004)** ya que existen propuestas de trabajos basados más bien en la experiencia que en la teoría. Por ejemplo, es una práctica de los investigadores orientados a casos, que a menudo buscan identificar comunalidades (*commonalities*) a través de un conjunto de casos, mientras se enfoca en un relativamente pequeño número de casos seleccionados a propósito **(Vaughan, 1986).** La búsqueda de comunalidades son la base de búsqueda-encuentro de conexiones empíricas **(Ragin, 2008).** Por ejemplo, en el caso de la administración del conocimiento del consumidor, la información del consumidor, acerca del consumidor y con el consumidor (como co-creador de valor) son indicadores de las pymes del sector de software en la zona metropolitana de la ciudad de Guadalajara, Jalisco, México, de logro como mejores prácticas de atención al mismo **(Mejía-Trejo et al., 2016).**

La evidencia indica que existe una relación basada en el modelo teórico propuesto entre la administración del conocimiento del consumidor y sus diversos indicadores así que es razonable especular en que el conjunto de los indicadores que describen al consumidor es un subconjunto del conjunto que establece la administración del conocimiento del consumidor en el sector de software en la zona metropolitana de la ciudad de Guadalajara, Jalisco, México y que existe un importante enlace causal entre ambos. Las comunalidades, permitirán explicar dicha relación en otros sectores de la industria.

Estrategias de análisis de las comunalidades

En general, se presentan dos estrategias de análisis, que se involucran en la búsqueda de comunalidades. Ver **Tabla 1.5.**

Tabla 1.5. Estrategias de análisis en la búsqueda de comunalidades

Estrategia	Descripción
1	La primera estrategia es examinar casos que comparten un resultado dado (por ejemplo, CEOs de pymes innovadoras recién creadas) e intentar identificar sus *condiciones causales* compartidas (por ejemplo, la existencia de un modelo de negocio o el cálculo correcto de precio a su producto/servicio). El término *condición causal* se usa para referirse a un aspecto de un caso que sea relevante de alguna manera para la cuenta del investigador o explicación de algún resultado. Es decir, *identificar las* condiciones causales *compartidas por casos con el* mismo resultado, es muy apropiado para la evaluación de condiciones necesarias.
2	La segunda estrategia, consiste en examinar casos que comparten una *condición causal* específica o, más comúnmente, una condición específica como combinación de *condiciones causales*, y evaluar si estos casos exhiben el mismo resultado (por ejemplo, los países que combinan regulación sobre prácticas de CEOs de pymes innovadoras y un bajo nivel de desarrollo económico ¿todos sufren de baja competitividad en sus resultados? En otras palabras, examinar casos con las mismas *condiciones causales* para ver si también, comparten el mismo resultado, es muy apropiado para la evaluación de condiciones suficientes, especialmente, combinaciones suficientes de las condiciones. Establecer las condiciones que son necesarias o suficientes es de gran interés para los investigadores que trabajan los aspectos macro políticos y/o macrosociales.

Fuente: Ragin (2008) con adaptación propia.

Como se observa, ambas estrategias son basadas en modelos teóricos. La primera estrategia es un examen de si las instancias de un resultado específico constituyen un subconjunto de instancias de una causa. La segunda estrategia, es un examen de si las instancias de una *condición causal* específica o la combinación de *condiciones causales* constituyen un subconjunto de instancias de un resultado. Ver **Figura 1.1.**

En dicha **Figura 1.1,** se presentan ambas estrategias son métodos para establecer conexiones explícitas. Por ejemplo, si es determinado que todos las pymes innovadoras de reciente creación tienden a ser exitosas debido a los estímulos de gobierno, entonces se ha establecido una conexión explícita *pyme innovadora* de reciente creación con estímulos de gobierno.

Figura 1.1. Diagramas de Venn
de dos diferentes tipos de investigación orientada a caso

Un conjunto de casos compartiendo una *condición causal*

Identificación de *condiciones causales* compartidas por instancias de un sólo resultado

Un conjunto de casos con el mismo resultado

Un conjunto de casos con el mismo resultado

Evaluación de si los casos con las mismas *condiciones causales* comparten el mismo resultado

Un conjunto de casos compratiendo *condiciones causales*

Fuente: Ragin (2008) con adaptación propia.

Probablemente, se ha encontrado que todas las pymes innovadoras de reciente creación que compartan un ambiente de bajo desarrollo económico con estímulos de gobierno pero un CEO con débil liderazgo fallan en su competitividad, de ser así, se establece una conexión explícita entre la combinación de condiciones previa y la falla en su competitividad. Como se observa, establecer conexiones explícitas no es lo mismo que establecer correlaciones. Por ejemplo, suponga que el radio de supervivencia a los tres años de pymes innovadoras de reciente creación con estímulos de gobierno y la falta de un plan de negocios no son exitosas, es del 70%, mientras el radio de supervivencia a los tres años de pymes innovadoras de

reciente creación con estímulos de gobierno y un cálculo erróneo de precio en producto/servicio no son exitosas, es del 25%. Claramente se observa que existe una correlación entre estos aspectos concebidos como variables (pymes innovadoras de reciente creación vs. estímulos de gobierno; plan de negocios vs. nivel de éxito; un cálculo de precio en producto/servicio vs. nivel de éxito). Sin embargo, ¿qué sucedería si la evidencia no es cercana a la aproximación del modelo teórico? En este caso, la evidencia mostraría una conexión correlacional, pero no una conexión explícita entre las pymes innovadoras de reciente creación y su nivel de éxito.

Cabe señalar, la importancia del uso de métodos basados en un marco teórico de la investigación, para establecer conexiones explícitas, los cuales no implican necesariamente el uso de los conceptos o el lenguaje de causalidad como el de *necesidad y suficiencia* **(Ragin, 2008).** Un investigador podría observar, por ejemplo, que en el caso un nivel alto de educación de un CEO de una pyme emprendedora tiene mayores oportunidades de ser competitiva sin prefigurar una conexión causal a dicha observación. La demostración de conexiones explícitas es fundamental para las ciencias sociales orientadas a la innovación ya sea si existe o no interés en causalidad necesaria o suficiente o cualquier otro tipo de causalidad.

Las correlaciones y las conexiones explícitas

La falta de coincidencia entre los métodos correlacionales y los estudios de conexiones explícitas son claramente visibles en la forma más simple de análisis orientado a variables, como la **tabulación cruzada 2 × 2 (Mejía-Trejo, 2019a,** en el Capítulo 9) de la presencia / ausencia de un resultado vs. la presencia / ausencia de una causa hipotética, como se ilustra en la **Tabla 1.6.**

Tabla 1.6. Cruce tabular de un resultado vs. *una condición causal*

Estatus	Condición causal (ausencia)	Condición causal (presencia)
Resultados (presencia)	**C1:** los casos aquí socavan el argumento del investigador	**C2:** los casos aquí apoyan el argumento del investigador
Resultados (ausencia)	**C3:** los casos aquí apoyan el argumento del investigador	**C4:** los casos aquí socavan el argumento del investigador

Fuente: Ragin (2008) con adaptación propia.

La correlación se enfoca simultánea y equivalentemente en el grado en el cual, las instancias de las causas producen instancias del resultado (el número de casos en la celda **C2** en relación con la suma de casos en las celdas **C2** y **C4**) y en qué medida, las instancias de la ausencia de la causa están relacionada con la ausencia del resultado (el número de casos en la celda **C3** en relación con la suma de casos en las celdas **C1** y **C3**).

El punto central es que la correlación es una técnica que recompensa a los investigadores por producir una gran cantidad de casos en celdas **C2** o **C3** y los penaliza por tratar casos en las celdas **C1** o **C4**. Por lo tanto, la herramienta de cruce-tabular, es una excelente opción para estudiar tendencias generales de casos cruzados. Sin embargo, un investigador de las ciencias sociales, orientado a la administración de la innovación, abocado en descubrir y analizar conexiones explícitas, sin embargo, estará interesado solo en componentes específicos de la información que estén agrupados y combinados en una correlación.

Por ejemplo, investigadores interesados en *condiciones causales* compartidas por instancias de un resultado, deberían centrarse en las celdas **C1** y **C2** de la **Tabla 1.5.** Su objetivo, sería identificar *condiciones causales* relevantes que depositan la menor cantidad de casos posibles (idealmente ninguno) en la celda **C1**.

Del mismo modo, los investigadores interesados en saber si los casos que son similares con respecto a las *condiciones causales* experimentan el mismo resultado, se centrarían en las celdas **C2** y **C4**. Su objetivo sería identificar combinaciones de *condiciones causales relevantes* que depo-

sitan tan pocos casos como sea posible (idealmente ninguno) en la celda **C4**. De estos ejemplos, está claro que la correlación tiene dos deficiencias principales, cuando se ve desde la perspectiva de conexiones explícitas:

1. Atiende solo a diferencias relativas, y
2. Combina diferentes tipos de evaluación causal.

Observe además, que una celda que es muy importante en el análisis correlacional, la celda **C3**, donde no está presente ni la causa ni el resultado, no es directamente relevante para la evaluación de cualquiera de los dos tipos de conexiones explícitas.

Por lo tanto, el estudio de conexiones explícitas implica una *descomposición* de la unidad más básica del análisis orientado a variables dependientes: *la correlación*. Esta descomposición, hace posible emplear estrategias de investigación cualitativa que se establecen fundamentalmente, en la naturaleza de un marco teórico para:

1. Estudiar casos con el mismo resultado a fin de identificar sus características causalmente relevantes, y
2. Estudiar casos con la misma combinación de condiciones causalmente relevantes a fin de identificar si exhiben el mismo resultado.

Es importante señalar, que la correlación no es simplemente una técnica de análisis multivariante dependiente bivariada. De hecho, es la piedra angular de la mayoría de las formas convencionales de la investigación social con orientación a la administración de la innovación, basada en variables, incluyendo algunas de las más sofisticadas formas de análisis cuantitativo disponibles en la actualidad. Una matriz de correlaciones bivariada, junto con las técnicas basadas en medias y las desviaciones estándar de las variables incluidas en una matriz de correlación, es todo lo que se necesita para calcular un análisis de regresión compleja, un análisis factorial (exploratorio / confirmatorio) e incluso un modelo de análisis de ecuaciones estructurales. En esencia, estas técnicas ofrecen diversas formas de representar las

correlaciones bivariadas en una matriz y las diversas relaciones parciales (por ejemplo, el efecto neto de una variable independiente en una regresión múltiple) que se pueden construir usando fórmulas basadas en tres o más correlaciones bivariadas. Dada la alta confianza sobre la correlación bivariada como piedra angular del análisis empírico, estas técnicas sofisticadas cuantitativas evitan el estudio de conexiones, descritas al momento. Esta deficiencia fundamental subyacente de la correlación, es la raíz del rechazo de los métodos de correlación de muchos académicos, que realizan estudios cualitativos y orientados a casos investigación **(Ragin, 2008).**

Análisis cualitativo comparativo y su relación con las conexiones explícitas

El análisis cualitativo comparativo (**QCA.** *Qualitative Comparative Analysis*), en contraste con las técnicas multivariantes con enfoque dependiente y de tipo correlacional, se basa en la teoría de conjuntos y, por lo tanto, es ideal para analizar conexiones explícitas, como las que se bosquejan en la **Figura 1.1.** Una característica especialmente útil de **QCA** es su capacidad para analizar la *complejidad causal*, definida como una situación en la que produce un resultado de varias combinaciones diferentes de *condiciones causales*, es decir, de diferentes trayectorias causales. Un ejemplo es el caso de que un investigador tenga suficientes razones que sustentan diferentes trayectorias de causalidad para la consolidación de la innovación abierta.

Al examinar el desenvolvimiento de cada uno de sus resultados, como casos con diferentes *configuraciones* de condiciones de causalidad relevantes, es posible, con el uso de **QCA**, identificar las trayectorias causales decisivas y por lo tanto, descomponer la *complejidad causal.*

Las tablas de verdad son la herramienta base que permite realizar comparaciones estructuradas, con enfoque para analizar la *complejidad causal* utilizando el **QCA (George, 1979).** Las tablas de verdad enumeran las combinaciones lógicamente posibles de *condiciones causales* y el resultado empírico, asociado con cada configuración. De hecho, una novedad en este caso, es el aplicar dichas tablas de verdad con trayectorias de causalidad difusas, no dicotómicas de origen, lo cual se tratará más adelante.

Por ejemplo, en un estudio propuesto por **Mejía-Trejo et al. (2016)** basado en un marco teórico de la investigación, que argumenta que existen trayec-torias causales que componen a la administración del conocimiento del con- sumidor (**CKM**), siendo: **CKMADI**. **CKM** como impulsor de innovación; **CKMS**. **CKM** como soporte; **CKMOSK. CKM** y otros recursos de conoci-miento**. CKMSEP**. **CKM** en satisfacción, experiencia, y desempeño donde dichas trayectorias de causalidad se combinan para producir un proceso de Innovación (**INNOV**). Ver **Figura 1.2.**

Figura 1.2. Modelo de relación procesos de innovación vs. administración del conocimiento del consumidor

Fuente: Mejía-Trejo et al. (2016)

La **Tabla 1.7.** muestra la *tabla de verdad* que ilustra la operacionalización de hasta dieceiséis argumentos o combinaciones de las cuatro variables que componen a la administración del conocimiento del consumidor (**CKM**) y que impactan a los procesos de innovación (**INNOV**).

Tabla 1.7. Tabla de verdad que relaciona las variables de la administración del conocimiento del consumidor (CKMS) con los procesos de innovación (INNOV)

Combinación	INNOV	CKMADI	CKMS	CKMOSK	CKMSEP
1	—	0	0	0	0
2	0	0	0	0	1
3	0	0	0	1	0
4	—	0	0	1	1
5	0	0	1	0	0
6	0	0	1	0	1
7	—	0	1	1	0
8	1	0	1	1	1
9	0	1	0	0	0
10	—	1	0	0	1
11	—	1	0	1	0
12	—	1	0	1	1
13	1	1	1	0	0
14	1	1	1	0	1
15	—	1	1	1	0
16	1	1	1	1	1

Notas:
"—". Los resultados indican que no existe una comprobación empírica con las combinaciones listadas en el renglón.
1. Es Sí
0. Es No

Fuente: Mejía-Trejo et al. (2016).

Como se observa, cuatro *condiciones causales* producen dieciséis combinaciones de dichas condiciones lógicamente posibles. En análisis más complejos, las filas, que representan las combinaciones de *condiciones causales*, pueden ser bastante numerosos, porque el número de combinaciones causales es una exponencial del número de *condiciones causales* (número de combinaciones causales = $\mathbf{2}^{2}$, donde ***k*** es el número de *condiciones causales*).

El uso de las tablas permiten descomponer la *complejidad causal* (**Ragin, 1987; De Meur y Rihoux, 2002**); el punto esencial, es que la

tabla de verdad es la base de la elaboración y formalización de una de las estrategias analíticas clave de la investigación comparativa: examinar los casos que comparten combinaciones específicas de *condiciones causales* para verificar si comparten el mismo resultado. El objetivo principal del análisis de la *tabla de verdad* es el identificar conexiones explícitas entre las combinaciones de *condiciones causales* y sus resultados. Al enumerar las diferentes posibles combinaciones de *condiciones causales* de manera lógica, es posible evaluar no solo el ser suficiente de una trayectoria causal específica, sino también el ser suficiente de las otras combinaciones lógicamente posibles de *condiciones causales* que pueden construirse a partir de éstas. Ver **Tabla 1.7.**

Por ejemplo, si los casos con las cuatro variables experimentan innovación (**INNOV**) y los casos con tres de cuatro variables presentes (y una ausente), también lo experimentan, entonces el investigador puede concluir que la *condición causal* que varía entre estas dos combinaciones es *irrelevante* para la trayectoria causal final así que las variables clave para el resultado son los restantes tres. Diversas técnicas y procedimientos para realizar simplificación lógica de patrones en tablas de verdad se detallan en **Ragin (1987, 2000), De Meur y Rihoux (2002)** y **Rihoux y Ragin (2009)**.

A menudo, el cambio de una propuesta de trayectoria causal hipotética a través de una *tabla de verdad*, estimula una reformulación o una expansión de la misma trayectoria, basada en un examen de casos relevantes. Por ejemplo, supongamos que la *tabla de verdad* reveló *inconsistencia* sustancial en algunas filas, es decir, casos que fallaron en provocar innovación (**INNOV**), aparte de las varias que lo hicieron. Esta *inconsistencia* en los resultados, indica al investigador que se necesita un estudio en profundidad de los casos.

Al comparar los casos de las filas que fallaron al provocar innovación (**INNOV**) vs. los que sí, sería posible proponer una trayectoria de causalidad más concreta. Suponga que el resultado de la comparación revela que los casos que no lograron provocar innovación (**INNOV**) tuvieron una variable empírica, más que de marco teórico de la investigación, que no se

consideró de origen, fue: el uso de tecnología e-commerce. Esta variable extra podría agregarse a la trayectoria de causalidad, y la *tabla de verdad* podría entonces ser respetada con cinco variables causales (y por lo tanto, treinta y dos filas).

La tarea del refinamiento de la *tabla de verdad* es muy exigente, ya que requiere conocimiento profundo de cada uno de los casos y muchas iteraciones de revisión con el marco teórico para la construcción de tablas de verdad. De hecho, la realización y revisión de las tablas de verdad disciplinan el proceso de investigación, proporcionando un marco para comparar los casos como *configuraciones* de similitudes y diferencias al explorar patrones de *consistencia* e *inconsistencia* con respecto a los resultados del caso.

Evolución del QCA

A fin de ubicar al lector, sobre la importancia de esta técnica y sus implicaciones, nos remitiremos a citar brevemente su nacimiento, el paso intermedio a **QCA** y hacia el **fsQCA** en las ciencias sociales y más allá.

Nacimiento

El término inglés *fuzzy* es traducido habitualmente al francés por *flou*, al italiano por *sfumato*, y al español por *difuso* o *borroso*. Dicho término fue acuñado en inglés, por L. A. Zadeh (1921-2017) para designar una categoría en el ámbito de las Matemáticas, y más en concreto en una de sus ramas: la Teoría de Conjuntos. Zadeh, oriundo de Bakú, Azerbayán y emigrante a los EUA, era por entonces, profesor de Ingeniería Eléctrica y Ciencia de la Computación de la Universidad de California, en Berkeley. Como ingeniero, había desarrollado sus investigaciones previas a 1965, en el campo de la Teoría de Sistemas, buscando refinar conceptos tales como *el estático* y *el adaptativo*. Su teoría de los conjuntos difusos, tiene como objetivo

tratar lo difuso de manera sistemática, aunque no necesariamente cuantitativa. **Zadeh** (1965, pp. 338-353) busca:

> *"…un sistema que proporcione una vía natural para tratar los problemas en los que la fuente de imprecisión, es la ausencia de criterios claramente definidos de los tipos de pertenencia... ya que la mayoría de las veces, la clase de objetos con que nos encontramos en el mundo físico real no dispone de criterios definidos de pertenencia".*

En la teoría clásica de conjuntos (llamada también *teoría abstracta de conjuntos*) desarrollada por G. Cantor (1845-1918), un conjunto es *una colección de objetos que poseen algunas propiedades muy generales, pero no se toma en consideración nada acerca de la naturaleza de los objetos individuales*. Los objetos individuales de la colección, llamados elementos del conjunto, tienen garantizada, y por igual, su pertenencia al conjunto, por ejemplo, los días de la semana, los meses del año, los números primos, etcétera.

En la teoría de conjuntos desarrollada por **Zadeh (1965, 1972)**, un *conjunto difuso*, es *una clase de objetos respecto de la cual un objeto no requiere o bien pertenecer o bien no pertenecer de manera absoluta a esa clase*. Cada objeto puede tener grados intermedios de pertenencia en el intervalo **(0,1);** la transición de la pertenencia a la no pertenencia es gradual más bien que abrupta, por ejemplo: las empresas innovadoras, los CEOs emprendedores, la innovación disruptiva. Estos conjuntos vienen determinados por referencia a dominios específicos (*locales*). Uno de los motivos por los que **Zadeh (1965)**, sugiere el concepto de conjunto difuso, es la *inconsistencia del formalismo matemático o lógico que es empleado para describir fenómenos o relaciones mal-definidos, vagos o subjetivos*. La tesis de la que parte **Zadeh (1965, 1972)** es la siguiente: *los elementos clave en el pensamiento humano no son números sino rótulos (marcadores lingüísticos) de conjuntos difusos*, por ejemplo, las clases de objetos en los que la transición de la pertenencia a la no pertenencia, es gradual, más bien que abrupta. La innegable realidad de lo difuso en los procesos del pensa-

miento y del razonamiento humanos sugiere que buena parte de la lógica propia del razonamiento humano *no es la lógica clásica bivalente o incluso polivalente, sino una lógica con cuantificadores difusos, con factores difusos, con valores de verdad difuso y con reglas de inferencia difusas*. Otra de las motivaciones, es lo que él denomina *principio de incompatibilidad* **(Zadeh, 1975, pp. 199-249):**

> *"...en la medida en que crece la complejidad de un sistema, en esa misma medida disminuye nuestra capacidad para hacer precisos y aun significativos enunciados acerca de su conducta, hasta alcanzar un umbral, más allá del cual la precisión y la significación (o relevancia) resultan, casi siempre, características mutuamente excluyentes..."*.

La teoría de lo difuso, al tener relación con las Matemáticas repercute altamente en las ciencias *blandas* como la Psicología, la Sociología, la Economía y las ciencias de la Administración entre otras, que figuran de las ciencias sociales.

El paso intermedio: QCA

Charles C. Ragin, el estudioso norteamericano, en 1987 publicó *The Comparative Method* **(Ragin, 1987)** comunicando a un público más amplio el método denominado **QCA** (*Qualitative Comparative Analysis*). El subtítulo que Ragin dio a su libro, *Moving Beyond Qualitative and Quantitative Strategies*, indica una discusión metodológica más amplia sobre las perspectivas cualitativas y cuantitativas, con una notable influencia en los años sucesivos. La introducción del **QCA (Ragin, 1987)**, contribuyó esencialmente al desarrollo de la metodología de las ciencias sociales por al menos cuatro razones **(Schneider y Wagemann, 2012)** ya que la obra de **Ragin (1987)**:

1. Ofrece a la investigación comparada, un método sistemático, riguroso y fundamentado en las matemáticas (álgebra Booleana y álgebra *fuzzy*) y

en la lógica formal la cual, hacía falta. La álgebra y los algoritmos propuestos, inmediatamente llevaron a usar, a los estudiosos, técnicas estandarizadas y bien definidas incluso en la investigación no estadística.

2. Propone a **QCA,** como una técnica que permite analizar, incluso un mediano número de casos, pues la numerosidad habría sido muy alta para el uso de técnicas empleadas en los estudios de caso, pero al mismo tiempo, muy baja para desarrollar un análisis estadístico ya que el número ideal para el análisis estadístico, es un mínimo de 30 casos, ofreciéndose más que esto y haciendo explícito, el aspecto del número medio de casos **(Ragin, 2000, p. 28).**
3. El **QCA** como método, permite que *el número de variables sea superior al número de casos*, cuestión que es considerada un vínculo relevante en la mayor parte de las investigaciones comparadas. No se menciona explícitamente dicho aspecto en sus publicaciones, dado que tal situación expone el problema de la denominada *diversidad empírica limitada* **(Ragin, 1987, p. 104)** que hasta el momento, ha puesto muchos obstáculos al desarrollo del **QCA**.
4. Presenta al **QCA,** como una estrategia para analizar hipótesis, basadas en *relaciones de teoría de conjuntos* (*set-theoretic relations*, con *hipótesis del tipo si… entonces…*). Este aspecto se resalta en las publicaciones de **Ragin de 2003 a 2006,** convirtiéndose en uno de los puntos más importantes para desarrollar.

Según **Ragin (2008, p. 183):**

> *"El objetivo de* **QCA,** *es derivar una lógica simplificada declaración que describe las diferentes combinaciones de condiciones vinculadas a un resultado*. Cada combinación de *condiciones causales,* que producen el mismo resultado, se denomina en ocasiones *tipo* o *tipología de configuración"* **(Fiss, 2010); Rihoux y Ragin (2009, pp. 33 y 66)** afirmando además:

*"El análisis cualitativo comparativo de conjunto de datos nítidos (***csQCA.*** Crisp Set Qualitative Comparative Analysis), fue la primera técnica de* ***QCA****, desarrollada a finales de 1980, por el Dr. Charles Ragin y el programador Kriss Drass. La investigación de Ragin, en el campo de la historia de la sociología, lo llevó a buscar herramientas para el tratamiento de conjuntos complejos de datos binarios, que no existían en la literatura estadística general. El Dr. Ragin adaptó a sus investigaciones, los conceptos de algoritmos Booleanos, que habían sido desarrollados en la década de 1950 por la ingeniería eléctrica, para simplificar los circuitos de conmutación a partir de los trabajos de* ***Quine (1952)*** *y* ***McCluskey (1966)****. En estos llamados algoritmos de minimización, se encontró un instrumento para identificar patrones de causalidad por conjuntura múltiple y una herramienta para simplificar datos de estructuras complejas de manera lógica y holística* ***Ragin (1987, p. viii)****. ...csQCA se basa en álgebra booleana, que usa solo datos binarios (0 o 1), y se basa en algunas operaciones lógicas simples [unión, intersección y complemento]. …En* ***csQCA****, es importante seguir una secuencia de pasos, desde la construcción de un tabla de datos binarios, a las fórmulas mínimas finales ... Se encuentran dos desafíos clave en esta secuencia, antes de ejecutar los procedimientos de minimización, que son: (1) implementar una dicotomización útil y significativa de cada variable, y (2) obtener una tabla de verdad (tabla de configuración) que esté libre de configuraciones contradictorias"... El procedimiento clave de la* ***csQCA*** *es la minimización booleana.*

Ragin (1987) extendió los alcances del **csQCA** a los conjuntos difusos **fsQCA**, porque razonó de que, categorizar las causas y los efectos de la ciencias sociales, en blanco y negro, no es realista. Los conjuntos de *datos difusos*, permiten evitar esto. De acuerdo con **(Rihoux y Ragin, 2009, p. 120):**

*"…**fsQCA** conserva aspectos clave del enfoque general de **QCA**, al tiempo que permite el análisis de fenómenos que variar por nivel o grado. ... El procedimiento **fsQCA** ... proporciona un puente entre conjuntos difusos y análisis de tabla de verdad convencional mediante la construcción de una tabla de verdad booleana que resume los resultados de múltiples análisis de conjuntos difusos. ... Puntajes de membresía difusos (es decir, el grado variable en el que los casos pertenecen a conjuntos) combinar evaluaciones cualitativas y cuantitativas. ... La relación teórica del conjunto clave en el estudio de la causalidad. La complejidad es la relación del subconjunto; los casos pueden evaluarse con precisión en términos de su grado de consistencia [subsecuencia] con la relación de subconjunto, generalmente con el objetivo de establecer que una combinación de condiciones es suficiente para un resultado dado".*

Tanto **csQCA** como **fsQCA,** son métodos de la teoría de conjuntos que permiten analizar, el marco teórico de una investigación. Se diferencian de los método cuantitativos convencionales métodos basados en variables (por ejemplo, la correlación y la regresión) en el sentido que expone **Fiss (2010):**

"…No desagregan los casos en aspectos independientes, como analíticamente separados, sino que en su lugar, tratan las configuraciones como diferentes tipos de casos".

Desafortunadamente aún se percibe cierto desconocimiento del **QCA**, principalmente en ciencias como las de la administración de la innovación. ¿Es solo un neologismo para denominar ciertas estrategias metodológicas que existían ya hace tiempo, o se trata en verdad de una solución para varios de los problemas metodológicos que enfrenta la investigación comparada? Se destaca que la discusión del mismo, aún está muy limitada en las aplicaciones principalmente de México y América Latina a pesar de su difusión extensa en Norteamérica, Europa y el Extremo Oriente.

Técnicamente, no es correcto hablar de un solo **QCA** porque detrás de la misma, se encuentra toda una familia de métodos. El principal promotor del **QCA,** el Dr. Charles Ragin, ha usado el término de *métodos comparativos configuracionales* (**CCM.** *Configurational Comparative Methods*) (**Rihoux y Ragin, 2009**) el cual, incluye tanto las diversas técnicas del **QCA**, así como las perspectivas relacionada con los sistemas *casos más semejantes con resultados diferentes* (**MSDO**. *Most Similar cases with Difierent Outcomes*) y el *diseño de sistemas más similares* (**MSSD**. *Most Similar Systems Design*) (**Morlino, 2005, p. 47; Przeworksi y Teune, 1970**). En el centro de los métodos del **QCA,** se encuentra la cuestión de las *condiciones suficientes* y/o *necesarias* (*o de combinaciones de condiciones que satisfagan tales características*) para un resultado (*outcome*), es decir, el **QCA** hace explícitas las relaciones entre las condiciones y el resultado **(Schnaider y Wagemann, 2012).**

La versión original del **QCA** es el denominado análisis cualitativo comparativo de datos nítidos (**csQCA**. *crisp-set* **QCA**) nombrado así solo después de la introducción de las otras variantes. El **csQCA**, se introdujo en 1987 en el libro *The Comparative Method* **(Ragin, 1987)** ganador del prestigioso premio *Stein Rokkan,* como un trabajo muy importante para la comparación en las ciencias sociales. En el **csQCA**, *todas las condiciones hipotéticas y el resultado deben ser forzosamente dicotómicas*. Por lo tanto, si queremos establecer por ejemplo, una relación causal entre empresas nacionales, con desarrollo tecnológico (*condición*) y su nivel de innovación (*resultado*) estamos obligados a dicotomizar los dos conceptos. Esta manera de proceder no siempre es satisfactoria, incluso no siempre es posible. Por otro lado, los investigadores regularmente no habían sido tan escrupulosos en la definición de las *dicotomías*. Por ejemplo, se puede observar que frecuentemente la mediana (el parámetro que corta una distribución en dos partes con idéntico número de partes) es usada para definir la *dicotomía* de un **csQCA (Schnaider y Wagemann, 2012)**. Esto consituye un error, pues se tiende a forzar la idea, por ejemplo, de que la mitad de las empresas nacionales, corresponde a ser innovadoras y la otra mitad, a no innovadoras independientemente del grado en el que pudieran encontrarse.

Incluso, aplicando formas de codificación más refinadas y más articuladas en aspectos teóricos, la necesidad de usar las dicotomías debilitaba considerablemente la aplicación del **csQCA**.

Más tarde, el Dr.Ragin respondió a las diversas críticas publicando un segundo libro titulado *Fuzzy-Set Social Science* **(Ragin, 2000)**, en el que propone utilizar los principios de los llamados *conjuntos difusos* (*fuzzy sets*) desarrollados en las áreas de la informática durante los años sesenta **(Zadeh, 1965; 1968)** e introducidos en los libros de matemáticas avanzadas **(Klir et al., 1997)**. Así, cabe preguntarse, ¿cuáles son las diferencias entre el conjunto de datos nítidos (*crisp sets*) y el conjunto de *datos difusos* (*fuzzy-sets*) y entre **csQCA** y **fsQCA** (*fuzzy set* **QCA**)? La diferencia estriba en dos aspectos y estos son, que en el **fsQCA (Schnaider y Wagemann, 2012):**

1. No es necesario dicotomizar el concepto que interesa analizar.
2. Tampoco es necesario, distinguir entre la presencia y la ausencia de un concepto (como pudiera ser la innovación,) como si es necesario hacerlo en el **csQCA**. Por el contrario, para cada caso se establece un *grado de pertenencia* al concepto (*membership*), **(Ragin, 2000, p. 3)**. Es decir para cada caso bajo análisis, se decide cuánto o en qué medida pertenece al concepto. De esta manera, puede señalarse que un país se considera innovador, pero no perfectamente innovador. Si fuera perfectamente innovador, recibiría un valor difuso (*fuzzy*) más alto, en relación a un país no perfectamente innovador. Así, el país no perfectamente innovador, recibiría un valor más alto respecto de un país no perfectamente no-innovador, que su vez pertenece más al concepto de la innovación, respecto de un país plenamente no-innovador. Para expresar determinado *grado de pertenencia,* se construyen las escalas difusas (*fuzzy-scales*) que atribuyen a los casos, valores entre **0** y **1**. Por ejemplo, un día de la semana como el jueves, pertenece al conjunto de los días, mientras que el mes de octubre, no, de acuerdo al conjunto de datos nítidos; sin embargo, en el conjunto de *datos difusos,* se hablaría de un *nivel de pertenencia al conjunto*. En nuestro caso de las empresas innovadoras, en el cual, todos los elementos (es decir, las empresas del

conjunto) pertenecen a cierto grado. Empresas más innovadoras pertenencen más respecto a aquellas empresas que no lo son; pero en cierta línea imaginaria (el punto de traslape, o de *crossover*), **(Ragin, 2000, p. 157)**, un elemento tiene mayor grado de pertenencia con un conjunto complementario **(Klir, et al., 1997, p. 74).** De esta manera se crean los conjuntos con *confines imprecisos* (*Ibídem,* p. 75). La propuesta de **Ragin (2000),** utiliza este principio del grado de pertenencia a un conjunto. Utilizando la terminología de las operaciones de los conjuntos difusos (*fuzzy-sets*) propone codificar (*calibrar*) también los conceptos de las ciencias sociales de una manera difusa, es decir, definir cuánto un caso (por ejemplo: país, empresa, organización, individuo, etc.) pertenece a un concepto (por ejemplo: innovación tecnológica, innovación por el modelo de negocios, pobreza, ingresos, religión, etcétera).

Sin embargo, debemos tener en cuenta que *no hay una diferencia sustancial entre los conjuntos nítidos y los conjuntos difusos* **(Schneider y Wagemann, 2012)**, como inicialmente se podría pensar: los conjuntos Booleanos no son más que conjuntos difusos, limitados a dos valores **extremos**, **(0)** y **(1)**. De esta manera, **fsQCA** es la versión más general y **csQCA** es el caso especial. Todas las reglas de la **fsQCA** también son válidas para **csQCA**. De lo anterior, se hace la firme recomendación de basarse en la definición de los valores difusos sobre todo (si es posible, e incluso exclusivamente) bajo parámetros teóricos **(Ragin, 2000, p. 150)**. De hecho, es altamente recomendable que el conjunto de datos difuso, sea visto, como una medida contínua y graduada, que ha sido altamente calibrada usando, tanto el conocimiento teórico como empírico sustantivo del investigador **(Ragin, 2000, p. 7; Ragin, 2007)**. Evidentemente es deseable la máxima transparencia sobre cómo se atribuyen los valores a cada caso pero también se aconseja **(Schneider y Wagemann, 2012)**:

1. Por un lado, que los valores difusos no pueden ni deben interpretarse como probabilidades, que manifiesten el concepto bajo análisis **(Cioffi-Revilla, 1981)**. Por ejemplo, imagine que se tienen dos vasos

con agua: uno de ellos, contiene agua tóxica con una probabilidad del **0.1**. El otro contiene agua, que pertenece al concepto de *agua tóxica* con un valor difuso o *fuzzy* del **0.1**. Si bebemos del primer vaso, no habrá consecuencias en el **90%** de los casos, mientras que el **10%** será letal. Por el contrario, beber del segundo vaso, provocaría seguramente malestares estomacales, aunque sea en cierta medida peligroso **(Schneider y Wagemann, 2012; Klir et al., 1997, p. 9)**.

2. Las escalas cuantitativas ya existentes, no deben ser traducidas proporcionalmente en valores difusos. **Ragin (2000, p. 167)** explica este procedimiento: por un lado, algunos conceptos demuestran solo un número limitado de valores que son teóricamente relevantes. Por otro lado, se corre el riesgo de que la evidencia empírica (es decir, nuestros casos) sean muy débiles como para permitir distinciones muy detalladas. Se permite utilizar también, el conocimiento que se obtiene por medio de la distribución descriptiva de variables cuantitativas; pero la *traspolación* en valores difusos, no debe ser automática. Con este propósito, debemos también considerar que, la definición de los valores difusos depende fuertemente del contexto, de tal manera, que los valores difusos son indicadores específicos para un sistema y no indicadores comunes **(Collier, 1998, p. 5). Ragin (2008, p. 85)** también ha introducido algunas formas semi-automáticas, más cuantitativas, de cómo definir los valores difusos.

El problema del tamaño N de la muestra

Según el propio **Ragin (2013, p. 171)** su objetivo inicial al escribir *The Comparative Method* era el de formalizar la *lógica del análisis comparado* tal y como estaba siendo aplicada por los investigadores comparatistas centrados en el análisis de casos **(Marx, et al., 2014)**, así como poner de manifiesto el fundamento relativo a la teoría de conjuntos que, a su juicio, subyacía en la lógica comparada. La idea inicial de **QCA (Rihoux y Marx, 2013, p. 166)**, es la de desarrollar una técnica comparativa para estudios de **N** pequeña o mediana que, a su vez, fuera capaz de relacionar el análisis

complejo de los casos con un grado satisfactorio de *parsimonia,* **(Marx et al., 2014)**. Es muy conocida la problemática de los estudios entre, aproximadamente, **5** y **50 casos**, en los que el objeto de estudio es demasiado reducido para el uso de las técnicas cuantitativas, pero demasiado amplio para elaborar un análisis profundo de cada uno de los casos. Varias investigaciones han encontrado que, en el **QCA,** la herramienta para colmar esta laguna metodológica, ya que permite el abordaje eficaz de la causalidad en las ciencias sociales orientadas a la administración de la innovación, cuando se trata de explicar la variación en un resultado de interés.

Es cierto que las técnicas **QCA** no son hoy en día dominantes en el ámbito de las ciencias sociales y mucho menos a las orientadas a la innovación, pese a su creciente importancia. Es frecuente encontrar dos aproximaciones estrictas; la *cualitativa*, basada en el estudio en profundidad de un caso o un número **N** reducido de ellos y la *cuantitativa,* basada en el estudio de un número **N** de casos muy elevado, normalmente mediante el uso de técnicas estadísticas. En este contexto **QCA** pugna por encontrar su propio espacio, si bien el uso de técnicas **QCA**, no está restringido a estudios de **N** mediana (**Berg-Schlosser et al., 2012**). Ver **Tabla 1.8.**

Tabla 1.8. Comparación entre aproximaciones en ciencias sociales

Aproximación	Número de casos (N)	Enfoque de análisis	Ejemplos de objeto de estudio
Cualitativa	N pequeña	Características del caso	La comparación de innovación entre dos empresas transnacionales
Configuracional	N mediana 8 entre 5 y 50 casos), aunque crecientemente aplicada a N grande	Interacción entre condiciones para producir un resultado. Narrativas acerca de los casos	La adopción o no, de una determinada política de innovación para todos los estados de un país
Cuantitativa	N grande	Impacto de variables independientes en los factores dependientes	El comportamiento de la innovación tecnológica a nivel mundial

Fuente: Berg-Schlosser et al., 2014 con adaptación propia.

Dicho espacio, sería tender un puente entre la investigación *cualitativa* y *cuantitativa* **(Grofman y Schneider, 2009, p. 662)** enfocándose en el análisis de los casos (*case-oriented approach*) *y no tanto en el impacto de las variables*. Es de reconocer que este temperamento epistemológico que opera en la **QCA** tiende a provocar cierto escepticismo entre los investigadores acostumbrados a otras metodologías y, por tanto, habituados a otros tipos de *presupuestos epistemológico.*

Debate del QCA

El debate entre detractores y defensores de las técnicas **QCA** es muy antiguo, teniendo como antecedentes:

- **Lieberson (1991, pp. 309-310),** defiende la *superioridad de la lógica probabilística vs. la lógica determinista,* con la que opera **QCA** ya que aquella permite minimizar las consecuencias de los errores de medición y sortear la imposibilidad de tomar en consideración todos los factores que determinan a un resultado, así como el impacto del *azar* en el fenómeno estudiado.
- Además, dicho autor critica la *lógica determinista* por el riesgo de rechazar una teoría o hipótesis con tan solo un único caso que la refute, sin tener en cuenta tendencias generales que se dan en el conjunto de casos.
- **Goldthorpe (1997)** sugiere algunas críticas más:
  a. QCA y su *lógica determinista* asume precipitadamente que la investigación científico-social puede obtener información cualitativa de un alto nivel *completa, de calidad, completa y ausente de errores.*
  b. QCA *no es capaz de integrar satisfactoriamente, el efecto potencial de variables* no tomadas en consideración o de errores en los datos usados. A este respecto, se debe recordar que es un problema común a otras metodologías, pero **QCA** establece una mayor exigencia en cuanto a la *transparencia, rigor y justificación* de la selección de condiciones **(De Meur et al., 2009, p. 158)**. **QCA,** además obliga a un proceso de diálogo permanente entre los casos y la teoría que permite comprender mejor las relaciones causales **(De Meur et al., 2009, p. 159)**.
  c. QCA *tiene un excesivo grado de sensibilidad* a la manera en que cada caso queda codificado para cada una de las condiciones. Además, se revela también contra lo que se considera un supuesto de los *análisis configuracionales*, que es la *independencia entre los casos*, que no sería capaz de dar cuenta de la historicidad de los procesos y la influencia de casos anteriores en caos posteriores.

d. Así también, acusa de que los *análisis configuracionales,* reposan excesivamente sobre el análisis inductivo de *configuraciones de condiciones*, esto es, las técnicas *configuracionales* a menudo no parten de una teoría aprobar sino construyen explicaciones *ad hoc*, a la luz de los casos deforma que cuando una teoría construida no es capaz de explicar un nuevo caso, se alude a la interrupción de la *homogeneidad causal* y *la necesidad de reelaborar la teoría*. En contrapartida, **Rihoux (2006, p. 682)** afirma que la *sensibilidad de los casos* es un problema desde el punto de vista de la *investigación cuantitativa* y que, desde la *perspectiva cualitativa,* basada en los casos, se trata en realidad de una fortaleza. **De Meur et al. (2009, p. 155)** afirman que la orientación hacia los casos permite que **QCA** *pueda y deba explicar todos los casos incluidos en el análisis,* evitando la exclusión de lo que en los estudios estadísticos se considerarían datos atípicos (*outliers*), lo cual, es particularmente útil en estudios de **N** pequeña y mediana. Para hacer frente a esta cuestión **QCA** ha desarrollado estándares exigentes de buenas prácticas relativos a *la homogeneidad de los casos, la diversidad de las configuraciones.*

e. **QCA**, frente a la crítica de la *dicotomización de los datos* en las primeras modalidades, no solo recuerdan que se han desarrollado modalidades *multicotómicas y difusas,* lo cual hace a la crítica obsoleta, sino también destacan que la *dicotomización* puede ser una valiosa herramienta para simplificar la **realidad (De Meur, et al., 2009, p. 148)**, que cabe añadir, es común a otras metodologías de las ciencias sociales y con oportunidad de aplicar a las orientadas a la administración de la innovación (por ejemplo: la más básica, el género del CEO de una empresa).

f. **QCA,** frente a la crítica al uso de *contrafactuales* (*o residuales lógicos o counter factuals o logical reminders*), los autores destacan su importancia para poder abordar poblaciones muy reducidas de casos y el hecho de que **QCA** obliga a hacerlos explícitos y justificarlos, mientras que, en otras metodologías, lejos de estar ausentes,

se hace un uso implícito y no justificado de los mismos **(De Meur et al., 2009, p. 152)**.

g. **QCA** no obliga a utilizar contrafactuales, pese a que en muchas investigaciones se usen *y que es posible realizar análisis excluyendo todos los contrafactuales*, mediante el uso de la *solución compleja.*

h. **QCA**, frente a la crítica de sus problemas ante la *temporalidad,* **De Meur et al., (2009, p. 161)** reconocen que este sigue siendo uno de los grandes desafíos de **QCA**, pese a que existen en marcha iniciativas para tratar de solucionarlo. A este respecto, existe una modalidad de **QCA** que tiene en cuenta la sucesión temporal de factores en la ocurrencia de un resultado, conocida como **TQCA** (*Tempora*l **QCA**) *que permite el análisis secuencial de las condiciones causales.* También existen propuestas para incorporar la *lógica panel* a **QCA** para comparar en varios puntos temporales (***t***) posibles variaciones en las soluciones a un fenómeno.

En todo caso, mientras el enriquecedor debate continúa vigente en las principales revistas sobre metodología en ciencias sociales, en términos empíricos, la utilización de técnicas **QCA**, dista de parar creciendo en los últimos años, particularmente desde el año 2002 a la fecha. Actualmente, existe un proyecto de la red internacional de investigadores que trabaja con **QCA** (www.compasss.org), la cual pretende crear el listado completo de publicaciones (artículos, capítulos de libro, documentos de trabajo, etcétera) que hayan aplicado algunas de las técnicas **QCA** en cualquier disciplina científica. Además, **QCA** habría extendido su uso más allá del análisis de matrices de datos **N** *mediana* empezando de forma modesta a utilizarse también para el análisis de matrices de datos de **N** *grande* **(Grofman y Schneider, 2009, p. 663)**. **Rihoux et al. (2013)** hablan así de un incipiente proceso de normalización de las técnicas **QCA**, cuya futura evolución no se atreven a predecir, pero que de momento se manifiesta no solo cuantitativamente sino también cualitativamente, en su aceptación de algunas de las revistas más relevantes indizadas en **Web of Science** o **SCOPUS** entre otros y localizables a través del *Journal Citation Report de Thomson* y/o *SCImago.*

La causalidad en QCA

El **QCA** propone examinar las llamadas, es decir, las relaciones entre conjuntos teóricos (*set-theoretic relations*). Un ejemplo de una relación de este tipo, es el análisis de las *condiciones suficientes y necesarias*. Existen numerosos ejemplos de hipótesis con estos principios; en sentido estricto, la hipótesis según la cual, el desarrollo económico avanzado produce innovaciones igualmente avanzadas, es una de éstas. Dos conceptos (desarrollo económico avanzado e innovaciones avanzadas) se interconectan a través de una relación de conjuntos teóricos (*si… entonces…*). En primer lugar, no nos interesa qué sucede si el desarrollo económico no es avanzado, ni qué causa la existencia de innovaciones no avanzadas o consolidadas. Por el contrario, nuestra hipótesis se refiere a la suficiencia de un desarrollo económico avanzado, para un nivel de innovación avanzado consolidado. Es decir, se tiene mucho por hacer con hipótesis sobre condiciones suficientes, incluso más de lo que se pueda pensar **(Ragin, 2000; Seawright, 2002, p. 180**; véase también el prefacio de **Ragin** al libro de **Schneider y Wagemann, 2007a, p. 13)**. De hecho, las hipótesis sobre condiciones necesarias, están muy difundidas en las ciencias sociales (véase la impresionante lista elaborada por **Goertz, 2003, p. 76)**. Además de la identificación de las *condiciones suficientes y necesarias*, el **QCA** va más allá; por ejemplo, es posible descubrir las denominadas *condiciones* **INUS** (*insufficient but necessary part of a condition which is itself unnecessary but sufficient for the result*, **Mackie, 1974, p. 62; Goertz, 2003, p. 68)**.

Si el resultado de un análisis **QCA** es por ejemplo, que las condiciones suficientes para una *innovación avanzada* **Y**, son por un lado el *desarrollo económico avanzado* **A,** *empresa fuertemente dedicada a la alta tecnología* **B,** y de un *sistema de productos/servicios diversificados* **C**, entonces los factores **B** y **C,** representan las condiciones **INUS** (*lnsufficient but Necessary part of a condition which is itself Unnecessary but Sufficient for the result*). Ninguno de los dos es, por sí mismo una *condición insuficiente* (*insufficient*), pero am bos de ben es tar pr esentes (*necessary part of a condition*) para formar una condición que no es

necesaria, pero suficiente (*which is itself unnecessary but sufficient*). La notación formal sería:

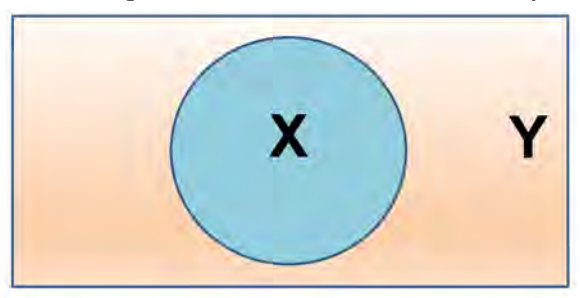
$$(A) + (B * C) \rightarrow (Y)$$

En esta fórmula:

1. El signo **(+)** corresponde al **OR** lógico; esta operación, prevee el valor difuso de **Y,** como el máximo de los valores difusos de **A** y de **(BC)**, de acuerdo a **Ragin (2000, p. 174)**. Ello significa que existe más de una condición suficiente (**A** o **BC**) para el resultado (vea *equifinalidad*. Nótese que la presencia de ambas condiciones suficientes, es una condición suficiente para el resultado).
2. De manera más específica, la parte (**BC**) de la ecuación, se escribe **(B * C)** donde el signo **(*)** corresponde al **AND** lógico; esta operación, prevee el valor difuso de **(B * C),** como el mínimo de los valores difusos de **(B y C),** de acuerdo a **Ragin (2000, p. 173)**.
3. La flecha **(→)**, señala que la fórmula es el resultado del examen de las *condiciones suficientes*. Convencionalmente, se utiliza una flecha invertida **(←)**, para indicar una *condición necesaria* y la flecha doble**, (← →),** para indicar que una *condición es necesaria y suficiente*.

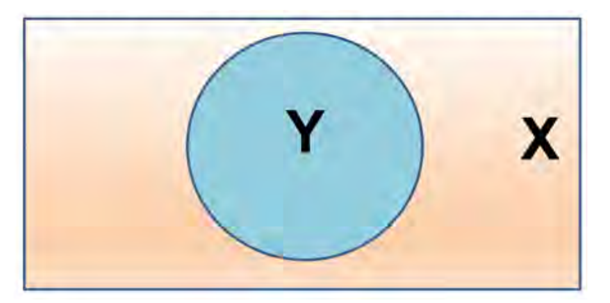
Por medio de la estadística, no es posible modelar estas relaciones. *Los métodos estadísticos, se basan en correlaciones entre dos variables* donde es posible establecer, si la variable dependiente aumenta o disminuye y cuánto, si varía el valor de la variable independiente, por sí sola o cuando interactúa con otras. Ello es muy útil (y en efecto, el **QCA** no logra informarnos sobre tales aspectos), pero las afirmaciones sobre la *suficiencia, necesidad o condiciones* **INUS** (*lnsufficient but Necessary part of a condition which is itself Unnecessary but Sufficient for the result*) no son fáciles de analizar con los métodos estadísticos. Entonces ¿por qué se habla de *métodos de teoría de conjuntos (set-theoretic methods),* si en realidad se trata del *análisis de condiciones*?, ¿qué relación existe entre las *condiciones* y los *conjuntos*? **Figuras 1.3 y 1.4.**

Figura 1.3. Diagrama de Venn de las *condiciones suficientes*



Fuente: Elaboración propia.

Figura 1.4. Diagrama de Venn de las *condiciones necesarias*



Fuente: Elaboración propia.

De ambas figuras se tiene **(Schneider y Wagemann, 2012)**:

a. La **Figura 1.3** representa una situación en la cual **X** es una *condición suficiente* de **Y.** Como se observa, allí donde exista **X** también estará **Y,** lo que significa que la presencia de **X** implica automáticamente la presencia de **Y,** pero no viceversa. En efecto, esto es la definición de una *condición suficiente*. Como se ve más adelante, el círculo de **Y** es mucho más grande que el de **X.** Esto nos sugiere que existen otras condiciones suficientes que pueden, en alternativa a **X,** explicar **Y.** La **Figura 1.3** representa también los dos conjuntos, de forma tal que uno de los conjunto, el **X,** es un *subconjunto* (*sub-set*) del otro conjunto, el **Y.** Es decir, cada elemento es parte del conjunto **X** que

es a su vez parte del conjunto **Y,** pero no todos los elementos de **Y** son parte de **X.** Para nuestra aplicación, ello significa que casi todos los casos (*elementos*), que tienen la condición suficiente, contienen también el *resultado*, pero no todos los casos que tienen dicho *resultado*, contienen, la condición suficiente (pues de otra forma sería una *condición necesaria*). La relación entre condiciones y resultado, es entonces una relación basada en subconjuntos, derivada de la más amplía teoría de conjuntos.

b. De esta forma se puede también interpretar la **Figura 1.4**, siendo **X** la condición necesaria para **Y, Y** no puede *existir*, sin **X.** Pero no siempre que exista **X** existirá también **Y.** Es decir, no es suficiente la existencia de **X** para que también haya **Y.** Esto es, **X** no es una condición suficiente. En este caso, **Y** es un subconjunto del conjunto definido por **X.** De la misma forma, se explican las direcciones de las flechas en las fórmulas. La dirección significa *implica*. Por lo tanto, la fórmula para la condición suficiente **X→Y** puede leerse como **X** implica **Y** (donde exista **X,** existirá **Y.** De esta manera fórmula, respecto de las condiciones necesarias es más clara: interpretando **X ← Y** como **Y** implica **X** (donde exista **Y,** existirá **X**), y se evitan las malinterpretaciones respecto de las direcciones de las flechas, como direcciones de mecanismo causal. Más claro, no es que el efecto **Y** cause la condición **X,** sino que la implica lógicamente.

No solo la centralidad de este tipo de relaciones derivadas de la teoría de conjuntos (*set-theoretic methods*) hace al **QCA,** particularmente diferente en relación con otras técnicas; resulta de hecho fundamental también la posibilidad de considerar la *complejidad causal* **(Braumoeller, 2003)** en la cual el **QCA** nos prepara.

De acuerdo a **Rihoux y Marx (2013, p. 168)**: **QCA** se define como un conjunto de técnicas *configuracionales.* Se trata de técnicas *holísticas,* en las que cada uno de los casos individuales es considerado como un todo complejo, es decir, como una configuración de condiciones (y resultados) que el investigador debe tomar en cuenta como tal en el transcurso de su

trabajo. Así, en lugar de descomponer su objeto de estudio en variables o factores, **QCA** propone no perder de vista los casos y volver a ellos con frecuencia, pues cada uno de ellos constituye una configuración de condiciones y de relaciones entre ellas, única, debiendo establecerse un diálogo entre teoría y casos **(Berg-Scholsser et al., 2009, pp. 6-7)**. De hecho, se considera **(Grofman y Schneider, 2009, p. 662)** que las técnicas **QCA** permiten entender mejor la complejidad causal en términos de: la necesidad y la suficiencia, el análisis configuracional, la equifinalidad, la causalidad por conjunción y la causalidad asimétrica, conceptos que no se encuentran en los analisis estadísticos, por lo que lo que distingue a **QCA** de las técnicas estadísticas no es el número de casos **(Grofman y Schneider, 2009, p. 666)**, sino los supuestos epistemológicos se desa-rrollarán brevemente y a continuación. Además, se debe recordar que el **QCA** es *determinista*, mientras que las técnicas estadísticas se basan en la lógica de la probalilidad. Esto es, mientras que la *lógica determinista* asume la ocrurrencia de un resultado dada la presencia de un factor, la *lógica probabilista* en sus aseveraciones causales, afirma que *la presencia de un factor incrementa la probabilidad o frecuencia de un resultado.* El carácter determinista de **QCA**, no es en sí un elemento negativo y de hecho en ocasiones puede ser considerado una ventaja, pero sin duda tiene fuertes implicaciones y sobre todo, obliga a que el investigador tome especiales medidas de rigor analítico. Una de ellas es cuidar por un control exhaustivo de la claridad de la matriz de datos, ya que un solo error de codificación en un caso puede conducir a resultados erróneos) y por maximizar la trans-parencia y proporcionar una detallada justificación en decisiones centrales de la investigación, como son la operacionalización de los fenómenos y la selección de las condiciones. Las técnicas configuracionales, por tanto, pueden ofrecer un interesante conjunto de herramientas a los investigadores dispuestos a ir más allá de las técnicas estadísticas clásicas o de las estrictamente cualitativas.

QCA y la complejidad causal

La literatura en **QCA** no utiliza los conceptos de *variables independientes* y *dependientes*. **QCA** se refiere a los conceptos de *condiciones explicativas* y *resultados* (*outcomes*). Esta distinción terminológica, tiene en realidad importantes *implicaciones teóricas, epistemológicas y metodológicas*. En lugar de estar basado en variables, como la *investigación cuantitativa*-o basado en los casos, como la *investigación cualitativa*, **QCA** se basa en la *diversidad* **(Gjolberg, 2007)**. **Ragin (2006)** ha enfatizado esta diferencia entre el *pensamiento configuracional* y el abordaje centrado en los *efectos netos*. En el caso de éste último, la investigación se centra en el *cálculo del impacto* de una *variable en otra con independencia del impacto de las demás variables*, que son mantenidas *constantes* **(Ragin, 2006, p. 15)**. El *análisis de efectos netos* aporta así una valoración cuantitativa, de la variación sobre la variable dependiente, que puede ser atribuida de forma individual, a cada una de las variables independientes. **Ragin (2006, p. 16)** sin embargo, considera que este tipo de investigación incurre en una serie de deficiencias, destacando el hecho de que cuando dos variables al estar fuertemente correlacionadas, por su *sobre determinación*, solamente podrán tener unos *efectos independientes netos modestos*.

La contraparte es que, en las ciencias sociales orientadas a la administración de la innovación, sin embargo, cuando los fenómenos se super-ponen, *tienden a reforzarse*, de forma que parece de alguna manera contra intuitiva que los científicos sociales confíen de forma casi exclusiva, en técnicas que favorecen la estimación de efectos separados, únicos y netos de cada variable causal **(Ragin, 2006, p. 17)**. El *pensamiento configura-cional* superaría este problema al centrarse en la *complejidad causal*. En el *pensamiento configuracional*, *los casos son entendidos como configura-ciones complejas de condiciones explicativas*, a la vez que una configura-ción puede ser entendida como una combinación específica de factores que produce un cierto *resultado (outcome)* **(Rihoux y Ragin, 2009, p. xix)**. Las *condiciones explicativas* no son estudiadas en **QCA** de forma independiente las unas de las otras, sino precisamente en su *yuxtaposición*, esto es, en su interacción tendiente a producir resultados (*outcomes*). **QCA** proporciona

así una *perspectiva holistica* sobre los casos, haciendo que se respete la complejidad y la singularidad propias de los mismos, a lo largo del análisis **(Berg-Schlosser et al., 2009, p. 6)**.

Como sugieren **Berg-Schlosser et al., (2009, p. 10)**, aunque **QCA** trata de alcanzar un nivel satisfactorio de *parsimonia*, también presta atención a la *complejidad causal*: en lugar de buscar la variable explicativa fundamental, **QCA** trata de identificar combinaciones de condiciones que conducen a un resultado, con una particular atención en cómo las condiciones se combinan y cómo y por qué las combinaciones conducen al resultado de interés. **Ragin (2006, p. 18)** considera que es precisamente al estudiar estas combinaciones cuando podemos entender las condiciones (*o contextos*) que permiten (*o impiden*) ciertas conexiones. La presencia o ausencia de ciertas *condiciones explicativas* o, mejor aún, las *combinaciones entre condiciones explicativas presentes o ausentes*, conducen a concretos resultados. Al mismo tiempo, la relación, en caso de existir, entre las *condiciones explicativas* y los resultados pueden ser de *necesidad o suficiencia*. No obstante, anotamos que:

a. Una condición se considera *necesaria* para un *resultado* si está siempre presente cuando el resultado ocurre, de forma que dicho resultado no puede producirse en ausencia de la condición.

b. Por el contrario, una condición se considera *suficiente* para un resultado si el resultado siempre ocurre cuando la condición está presente, incluso aunque este resultado pueda producirse también debido al impacto de otras condiciones explicativas.

QCA utiliza diferentes formas para expresar este tipo de soluciones (véase la **Tabla 1.9**).

Tabla 1.9. Notación clásica y notación moderna QCA

Aproximación	Presencia de la condición (A)	Ausencia de la condición (A)	Lógico AND (Conjunción)	Lógico OR (Unión)
Notación clásica	A	a	*	+
Notación Moderna	a	~a	*	+

Fuente: Medina et al. (2017).

La necesidad y la suficiencia

El enfoque causal, basado en las relaciones de la *necesidad* y la *suficiencia,* como se trató líneas arriba, hace que desde esta perspectiva, no estar enfrente a relaciones causales fruto de correlaciones, sino ante la posibilidad de que un factor sea necesario (pero no suficiente) o suficiente (pero no necesario) para otro, sin que sea descartable que la condición sea al mismo tiempo necesaria y suficiente para contribuir al resultado **(Berg-Scholsser et al., 2009, pp. 15-16)**. Las ideas de *necesidad y suficiencia* son capaces de ofrecer análisis más sofisticados y detallados de las relaciones causales y su naturaleza en cada caso, que son difíciles de capturar con las técnicas estadísticas ordinarias.

Análisis configuracional

Al ser combinaciones lógicas de *condiciones causales* y no variables independientes aisladas las que conducen al resultado **(Grofman y Schneider, 2009, p. 666)**, hace que en este tipo de investigaciones, tenga poco sentido tratar de estimar el *efecto neto* **(Mahoney y Goertz, 2006, p. 235)** de una condición individual sobre el resultado. El *análisis configuracional,* ofrece una ventaja comparativa respecto a las técnicas estadísticas clásicas, que tienen más dificultades de analizar de forma directa y comprensiva la forma en que diferentes factores interactúan entre sí en la producción de fenómenos.

La equifinalidad

La idea de la *equifinalidad* implica que diferentes combinaciones de factores puedan producir el mismo resultado **(Grofman y Schneider, 2009, p. 666).** Lo anterior conduce a un término complementario como el de *causación por conjuntura múltiple* **(Rihoux y Marx, 2013, p. 168)**, el cual implica:

a. *Es la combinación de condiciones la que produce un resultado.*
b. *Las diferentes combinaciones de condiciones pueden conducir a ese resultado.*
c. *Que una misma condición puede tener distintos impactos causales depediendo de las demás condiciones con las que interactúe.*

La *equifinalidad*, se expresa en que puede existir más de una condición suficiente (*pero no necesaria*) para producir un resultado. En el ejemplo de arriba:

$$\mathbf{(A) + (BC) \rightarrow (Y)}$$

La *solución equifinal* significa qué tanto de la condición **A** como de la condición (compuesta) **(B * C),** representan las *condiciones suficientes*. En otras palabras, incluso si **A** no estuviera presente, el resultado podría obtenerse, a través de la presencia de las condiciones alternativas, **(B * C),** y viceversa. Obviamente existe la posibilidad que ambas condiciones estén presentes, al mismo tiempo.

La diferencia con las ecuaciones de los métodos estadísticos (*por ejemplo, la recta de regresión*) es clara: en la regresión, las variables independientes no son alternativas, pues todas contribuyen en explicar las variaciones de la variable dependiente. Como afirma **Ragin (2003a, p. 8)**, esto puede significar que las variables están en cierta *competencia,* la una con la otra; cada una contribuye con cierto porcentaje. Si las variables independientes, por el contrario, fueran alternativas la una por la otra, la lógica de la regresión implicaría que solo una de esas *venciera,* o sería superior a la otra (*ibídem*). Así, puede suceder, que una variable que se explica como cualquier otra, sea excluida de una ecuación de regresión solo porque no contri-

buye tanto como las otras. Obviamente, esto no puede suceder si todas las variables independientes de la regresión son perfectamente independientes una de la otra.

Esta *competencia*, no existe en la equifinalidad del **QCA**. Por el contrario, las *condiciones causales* cuentan como *ingredientes* con los cuales una relación causal compleja puede ser modelada. Cabe decir, que alguna de las *condiciones causales* tienen la posibilidad de ser *superfluas*, aunque permanezcan siempre como condiciones suficientes. En otras palabras, la equifinalidad **(Schneider y Wagemann, 2012)**:

1. Se refiere sobre todo a las condiciones suficientes.
2. En relación a las condiciones necesarias, forzosamente *no pueden existir alternativas equifinales, porque si existieran las condiciones no serían entonces necesarias.*
3. La única manera en la cual la equifinalidad existe para las condiciones necesarias es por medio de las *equivalencias funcionales*.
4. Una afirmación equifinal (por ejemplo, **A + B**), se considera como una única condición necesaria. Pero debe decirse también que el desarrollo de estas reflexiones dentro de la **QCA** apenas está comenzando.
   Esto ha sido discutido desde hace algunos años y se le conoce como el fenómeno de las *condiciones* **SUIN** (*Sufficient, but Unnecessary part of a factor that is Insufficient, but Necessary for the result*) **(Mahoney y Barrenechea, 2017)**.

Lo anterior, parece corresponder a nuestras hipótesis en las ciencias de la administarción de la innovación comparadas, es decir, sabemos de ciertos resultados (o *outcomes*, como por ejemplo los grados de innovación de un país) no se logran de manera unívoca, o sea, con solo una posibilidad de introducir diversos factores a manera de una suma. Por el contrario, regularmente la realidad social, nos obliga a razonar de manera *equifinal* y a descubrir varias vías al resultado bajo análisis.

Por cierto, en adición al esquema *equifinal*, también el **QCA** tiene la capacidad de *evaluar la importancia matemática, de cada una de las condi-*

ciones equifinales. Las medidas de *cobertura* **(Ragin, 2006)** como instrumento de medición, no solo permiten evaluar cuánto, cada condición en lo particular, explican (*o cubren*), sino también cuánto es indispensable de cada una de ellas en lo individual, es decir, *cuánto explica que no haya sido explicado por otras condiciones.* Justamente, es a través del procedimiento de *implicaciones principales* (*prime implicants*) que técnicamente es posible integrar las *condiciones suficientes*, en una solución que sería superflua desde el punto de vista *parsimonioso*.

Causalidad por conjunción

Uno de los problemas centrales de la *metodología comparada* **(Morlino, 2005, p. 114)**, es la causalidad por conjunción, sobre todo en la aplicación de los cánones de Mill (**Lieberson, 1991;** para los cánones de Mill, cfr. **Mahoney, 2003; Morlino, 2005, p. 86; Skocpol, 1984, p. 379**). Lo anterior reside en el hecho de que en ocasiones, *una condición por sí sola no es suficiente, pero debe existir para ser combinada o en el contexto de más de una variable.* En nuestro ejemplo:

$$\mathbf{(A) + (BC) \rightarrow (Y)}$$

Como se ha observado, las condicones **INUS** (*Insufficient but Necessary part of a condition which is itself Unnecessary but Sufficient for the result*)**,** se basan en la idea de la causalidad por conjunción. Es un hecho, de que la estadística suele *despreciar* situaciones en las cuales, dos variables estén altamente correlacionadas traduciéndolas en *multicolinearidad*, por lo que se señala claramente, que la *multicolinearidad* y la *causalidad por cojunción* no son lo mismo, porque **(Schneider y Wagemann, 2012)**:

a. La *multicolinearidad*, se refiere a un aspecto descriptivo dentro de un conjunto de variables independientes.
b. La *causalidad por conjunción*, por el contrario, se refiere al momento causal de un análisis, mostrando que los factores individuales causales, no pueden existir por sí mismos.

No obstante, es cierto que con regularidad, la *multicolinearidad* se convierte en un problema del análisis estadístico por la razón de *que la estadística no tiene ningún instrumento para tener en cuenta la causalidad por conjunción*. Por lo que hay mucho por hacer en este campo, dado el mundo real, es de naturaleza multicolinear. Esto permite prever, que la mayor parte de las investigaciones por ejemplo, en la administarción de la innovación, producirán resultados con roles causales que corresponden a cada una de las condiciones, aunque también se producirá de la combinación de cada una de las mismas. Por ejemplo, consideremos una situación en la cual el análisis de las condiciones suficientes produce el resultado:

$$\mathbf{(AB) + (\sim AC) \rightarrow (Y)}$$

Donde también la **~** significa negación o complemento y también se representa haciendo la letra minúscula, tal que **(~A) = (1- A) = (a)**. Así, de la expresión original, se observa que la condición **A** tiene dos roles causales diferentes, las cuales dependen del contexto en el cual aparece:

a. En combinación con la condición **B, A** debe estar presente;
b. Por el contrario, en combinación con la condición **C, A** debe estar ausente **(~A = a).**

Como se sabe, a veces en las investigaciones, sucede en realidad que *una condición que es favorable para un resultado en ciertas circunstancias, no lo es en otras.* Incluso, *la condición puede ser desfavorable en otras circunstancias.* Un ejemplo, son ciertas estrategias para la detonar la actividad de la innovación que podrían funcionar en ciertas sociedades o ambientes, pero podrían ser contraproducentes en otras. La capacidad versátil del **QCA** sobre estos aspectos de la *causalidad por conjunción* significa que *no solo se habla de casos sino principalmente de configuraciones* **(Ragin, 2000, p. 64)**. Desde este punto de vista, cada caso viene como un rompecabezas, *despiezado* por así decir, en sus unidades constituyentes utilizando las *tablas de verdad* como su instrumento de análisis básico. Lo anterior, trae como consecuencia que se llega a examinar la estructura y

distribución de tales unidades constituyentes. Este procedimiento, permite observar y determinar que, *algunas combinaciones de unidades constituyentes simplemente no existen*. Este aspecto, conocido como *diversidad empírica limitada* o *diversidad limitada* **(Schneider y Wagemann, 2003)** sea un aspecto de gran relevancia en el estudio y análisis del **QCA**.

Causalidad asimétrica

De acuerdo con **Grofman y Schneider (2009, pp. 666-667)**, el énfasis en la causalidad asimétrica, según la cual la presencia de un fenómeno y su ausencia, requieren de análisis y explicaciones separadas. Como afirma **Ragin (2013, p. 173)**:

> *"...los coeficientes de correlación utilizados en las técnicas estadísticas, son simétricos, de forma que cuando se valora la conexión entre la presencia de una causa y un efecto, se está haciendo también una valoración acerca de la conexión entre la ausencia de la causa y el efecto. Las relaciones basadas en teoría de conjuntos, sin embargo, son asimétricas..."*.

Un breve ejemplo de esto, es la asociación de ideas de que las empresas desarrolladas son innovadoras tecnológicamente y no implica que las empresas no desarrolladas sean innovadoras en productos/servicios.

La *causalidad asimétrica*, significa que el conocimiento de las causas para un resultado no implica necesariamente que conozcamos igualmente las causas para un resultado contrario. Por ejemplo, si sabemos cuáles son los factores a través de los cuales podemos identificar un país con innovación no necesariamente sabemos cuáles son aquellos factores que generaran un sistema de país no innovador. Por el contrario, la interpretación de resultados estadísticos, regularmente se basa en que la explicación de un país no-innovador, podría derivarse directamente de una ecuación que explica la existencia de un país innovador. Un país no-innovador, en los métodos cuantitativos, es simplemente vista como la parte inferior de una escala para medir a los países innovadores.

Como el **QCA** muestra una relación de argumentos técnicos **(Schneider y Wagemann 2007a, p. 127)**, esta manera de proceder funciona si y solo si todas las *configuraciones*, es decir, todas las combinaciones de unidades constituyentes de los casos, existen en la realidad y son parte de nuestra base de datos; en otras palabras, sino existe *diversidad empírica limitada* o *diversidad limitada*. El razonamiento matemático de lo anterior, es que solo en aquel caso se aplica la Ley de DeMorgan **(Klir et al., 1997, p. 37)**, que permite volver a calcular las *condiciones causales* para el resultado, de forma que se entiendan las condiciones para el contario del mismo resultado.

Para lograr lo anterior, es absolutamente necesario en el ámbito del **QCA** se desarrollen siempre dos análisis: uno para el *resultado*, y otro para su *contrario* **(Schneider y Wagemann, 2007b)**. En el caso de la *diversidad empírica limitada* o *diversidad limitada* puede acontecer, que el tratamiento de las combinaciones faltantes en dos análisis, tanto del resultado como de su contrario, sean contradictorios **(Vanderbrought y Yamasaki, 2003; Schneider y Wagemann, 2012)**. Con la ayuda del software, las contradicciones pueden visualizarse, pero queda en los investigadores el problema de como resolverlos.

Con el **QCA,** como se puede observar, existe la posibilidad de considerar varios aspectos de la *complejidad causal*. Sin embargo, todas estas consideraciones traerán consigo que los análisis **QCA** se conviertan en algo muy complicado, teniendo en cuenta los aspectos hasta aquí presentados. Por otro lado, resulta positivo que se requiera, la máxima atención del investigador a sus casos, a sus hipótesis y sus teorías. De esta forma, en ocasiones el resultado de un **QCA** puede ser que las condiciones y los casos elegidos por el investigador, no se adapten o sean poco precisos para estudiar el fenómeno de interés.

Los métodos cuantitativos, han habituado a los investigadores a reducir sus hipótesis a preposiciones simplistas, porque ésta es la única manera para analizarlas con los métodos disponibles. En este sentido, **Wagemann y Memoli (2007)** hacen una demostración sobre el **QCA** que revela de manera muy concreta, que los *modelos parsimoniosos* aplicados en la esta-

dística *no llevan directamente, a una explicación completa de un resultado*. Es simplemente imposible, por ejemplo, querer explicar la innovación de un país, solo con factores económicos y con el sistema educativo, seleccionados bajo el postulado de la *parsimonia*. De hecho, con dicho postulado se corre el riesgo que las hipótesis en realidad se demeriten drásticamente **(Munck, 2001)**. De esta manera, **QCA** tiene la posibilidad de ayudar al investigador a darse cuenta de las estructuras complejas y a integrarlas a sus explicaciones.

Las comparaciones en el QCA

El **QCA** como metodología, se ha presentado como una novedad. Pero ¿qué tan nueva es? Y si es así, ¿en dónde radica su novedad? Los principios de la *complejidad causal* presentados al momento, son los fundamentos del análisis comparado, tal y como se ha desarrollado desde los tiempos de la filosofía Aristotélica. El **QCA,** es solo la etapa más reciente que se dirige hacia una mejor sistematización de los conceptos usados. Cabe anotar que, mientras Aristóteles realizaba un sistema de comparación que después fue copiada por Cicerón, o que, en tiempos más recientes, la comparación de Tocqueville, entre Estados Unidos y Francia, todos se basaban en una *metodología implícita*. Es John Stuart Mill quien se logra encontrar con un *sistema metodológico* a través de *los cánones,* **(Mahoney, 2003; Morlino, 2005, p. 86)**. Aunque con una historia de más de 150 años, su propuesta es una importante referencia vigente y regularmente utilizada. Cabe decir también, que la aplicación de su sistema a la realidad empírica, *no está exenta de problemas*. Se aprecian en primera instancia, *demasiado simples* para ser utilizados en un análisis sistémico de forma plena. La combinación de dichos métodos, *trae consigo una complejidad que corre el riesgo de requerir demasiada concentración analítica.*

El sistema de Mill, se basa implícitamente en una forma de álgebra que fue desarrollada más o menos en el mismo periodo, al álgebra Booleana, la cual tiene características que la hacen adaptable al *análisis comparado* respecto del *álgebra lineal*. **Ragin (1987)** logra combinar efi-

cientemente, la versión formalizada del álgebra Booleana con la metodología de Mill.

Por lo tanto, podemos afirmar que el **csQCA** (*la versión dicotómica)* no es otra cosa que una aplicación del método de Mill con las cuales tenemos la posibilidad de analizar sin mayores problemas, las situaciones más complejas. Desde este punto de vista, para la *metodología comparada,* **QCA** es un instrumento poderoso. Pero el aspecto *nuevo* del **csQCA** (y del **QCA** en general) es probablemente otro: el aspecto de la rigidez de la perspectiva formalizada del análisis comparado, hace estar atentos a la vista de los investigadores, de varios aspectos de sus análisis, que de otra manera no se observarían. Una forma de ilustrar esto es el fenómeno de la *diversidad empírica limitada* o *diversidad limitada*. **QCA** descompone los casos en *con iguraciones* con sus características. Por ejemplo, la empresa **A** tiene innovación tecnológica con un alto número de productos diversificados. La empresa **B,** es diferente a la empresa **A,** por tener un bajo número de productos diverisificados pero comparte con la empresa **A** el tener innovación tecnológica. La empresa **C,** es diferente a la empresa **B** por una cuestión (tiene bajos ingresos por innovación) y diferente de la empresa **A** por dos cuestiones (tiene bajos ingresos por innovación y el bajo número de productos diverisificados) En conjunto, los tres componentes (*dicotómicos*) de los sistemas de innovación podrían producir ocho *configuraciones*. Dicha cuestión está fuertemente relacionada, con la construcción de tipologías **(Morlino, 2005, p. 74)**.

Nótese que, para ver en realidad estas ocho *con iguraciones* necesita-ríamos de ocho casos individuales. Si consideramos que la empresa **D** es igual a **B,** usando las tres categorías: la empresa **D** tiene también innovación tecnológica con una baja cantidad de productos diversificados. Entonces, para tener las ocho *con iguraciones* son necesarios ocho empresas diferentes según nuestras categorías. Así, el *número de con iguraciones* posi-bles aumenta de manera exponencial. La fórmula es 2^k, siendo k el número de componentes de las *con iguraciones.* Por ejemplo, en el modelo de inno-vación por mercadotecnia digital, propuesto en **Mejía-Trejo (2017a)**, se tienen diez componentes a saber: misión-visión (**MVS**.*Mission-Vision);* propuesta de valor

(**VAL**.*Value Proposition*); Mercado (**MKT**.*Market*); Metas (**GST**. *Goal Settings*); Estrategia (**STG**.*Strategy*); Táctica (**TAC**.*Tactics);* herramientas de mercadotecnia (**DMT**. *Digital Marketing Tools)*; Planeación (**PLN**.*Planning*); Desempeño (**PER**.*Performance*); Ganancias (***PRO***.*Pro i-tability*), del que se lograrían al menos **2 a la 10**= **1,024** combinaciones dife-rentes, por lo tanto **1024** empresas por caracterizar.

Recordemos que los componentes de las *configuraciones*, son también al mismo tiempo, cada una de las condiciones que se analizan respecto de su *suficiencia* y su *necesidad* para un *resultado*. Este aspecto, nos señala dos puntos equivocados que a veces se mencionan sobre el **QCA (Schneider y Wagemann, 2012)**:

a. Por un lado, *el **QCA** no es un instrumento para el análisis de situaciones con **n** pequeñas. El número de casos no debe ser inferior (digamos entre ocho y diez aunque, ya con diez el análisis es complicado),* pues de otra manera, solo se podrían analizar un número muy bajo de condiciones.
b. Por otro lado, el **QCA** no resuelve el problema de *muchas variables, **n** limitado* **(King et al., 1994, p. 213.; Morlino, 2005, p. 47)**. De allí, que el problema de no tener todas las combinaciones posibles a disposición, puede tener efectos sobre el desarrollo de los análisis.

Se sabe también, que no es fácil (sino, imposible), tener todas las combinaciones posibles de los componentes de cada una de las *configuraciones*. En tal situación se habla de *diversidad empírica limitada*, o *diversidad limitada*, existiendo tres razones por las cuales el conjunto de casos no está completo **(Schneider y Wagemann, 2012)**:

1. *La configuración no puede existir lógicamente*. Se trata de los famosos (*como empresas sin personal*) los cuales, por razones que van más allá de las ciencias de la administración orientadas a la innovación, no encontraremos.

2. *La configuración puede existir lógicamente, pero no en la práctica*. Hasta ahora, falta empíricamente la configuración con las características **A**= *La persona, es el CEO de la empresa más innovadora en tecnología de los Estados Unidos de América* y **B**= *La persona es una mujer de origen latino.* Este caso, nos muestra que esta imposibilidad tiene tendencia de cambiar velozmente, no obstante en la actualidad fallaríamos al examinar el efecto conjunto de un CEO que es al mismo de la empresa más innovadora tecnológicamente de los Estados Unidos de América y sea mujer de origen latino. Comprendemos entonces que nuestra realidad empírica, es simplemente *limitada* pero que no debemos negar la imposibilidad de llegar a las diferencias máximas.
3. *La configuración existe empíricamente, pero o no la hemos encontrado o no la hemos incluido en nuestra selección de casos*. Esta posibilidad, obviamente, debe evitarse. En un **QCA** cabe señalar, que raramente se tiene la posibilidad de influenciar la selección de casos. Regularmente, se examinan poblaciones enteras como las empresas por regiones de un país o de varias regiones de todos los países, o clústers tecnológicos por ejemplo.

Pero entonces ¿cómo resolver el problema de los casos que empíricamente nos faltan? La respuesta inmediata es*, que no se tiene la posibilidad de resolver este dilema*, pero lo que sí es posible, es enfrentarlo a través de dos estrategias centrales **(Ragin, 1987, p. 104; Schneider y Wagemann, 2012)**:

- La primera posibilidad, se basa en la idea que, en consonancia con la *diversidad empírica limitada*, el *resultado* no es unívoco. Así, suponiendo que los casos faltantes existieran, el *resultado* puede variar fuertemente. Por ejemplo, si faltan tres de las *configuraciones*, entonces no menos de ocho resultados son posibles, uno para cada posibilidad de codificar el resultado de las configuraciones faltantes (**2^3**). La primer forma, para afrontar el problema de la *diversidad empírica limitada* puede parecer un tanto mecánica y banal: hacer que el software de la

computadora simule todos los *resultados* y eligir el resultado *más parsimonioso* (*most parsimonous solution*; en el lenguaje del **QCA**, se dice también de un resultado basado en *simplifying assumptions*, **Ragin, 1987, p. 113**). Esta estrategia parece demasiado ambigua y poco científica. No es razonable dejar a la computadora la discusión y análisis sobre cómo un CEO de la empresa más innovadora en tecnología de los Estados Unidos de América, siendo mujer latina actuaría en el mundo empresarial. Los resultados de este procedimiento, pueden parecer como altamente manipulables. No obstante, al final, la estrategia no es del todo inútil. Además de presentarnos uno de los resultados posibles (de hecho, no se sabe nunca cuál resultado es el *correcto*; puede ser el *resultado parsimonioso* como también los otros resultados). El *resultado más parsimonioso* nos indica una especie de solución mínima, una que no es posible simplificar. Se puede estar seguro, que las relaciones causales no son más simples del *resultado más parsimonioso.*

- La segunda estrategia, por el contrario es mucho más conservadora. Son tomadas en consideración solo aquellas configuraciones que realmente existen. Los resultados de las configuraciones faltantes por el contrario, son codificados todos con (**0**); o sea, la idea principal es que esos no tendrían un resultado si existieran. Es fundamental, para esta estrategia, no simular eventuales efectos positivos para los casos que en verdad no existen. Esta estrategia (*blanket assumptions*) parecería problemática por dos razones **(Schneider y Wagemann, 2012)**:
  - **a.** En primer lugar, los resultados se vuelven velozmente muy complejos (porque no existen tantas posibilidades de encontrar estructuras generales en los datos). Es equivocado decir que esta estrategia produce automáticamente el resultado más complicado de todos los resultados posibles. Los *blanket assumptions* producen el resultado más complicado solo en el caso en el cual las simulaciones que llevan a la *solución parsimoniosa* codifica todos los resultados de las *configuraciones faltantes* con (**1**) y regularmente no son más que descripciones de las *tablas de verdad.*

b. Otro problema, se relaciona con el principio de *causalidad asimétrica*. Estamos obligados, no solo a analizar las causas para el *resultado,* sino también para su *contrario*, el uso de los *blanket assumptions* producirá automáticamente *contradicciones*. Como se sabe, el *resultado de una configuración* que no existe es codificado como (**0**). Pero también, el contrario del resultado de la configuración que no existe debería ser codificado como (**0**) siguiendo la perspectiva de los *blanket assumptions*. En ambos análisis, la configuración no existe, ya que una vez que el resultado es codificado como (**0**) por lo tanto, su contrario sería (**1**), mientras en el otro, el contrario del resultado es codificado como (**0**) (y por lo tanto el resultado mismo sería **1**). Este problema ya se ha considerado en la literatura sobre el **QCA (Vanderbrought y Yamasaki, 2003)**.

La *diversidad empírica limitada* o *diversidad limitada*, es un problema fuerte en el **QCA**. No debe sorprender, que las soluciones propuestas para este dilema vayan más allá de las dos posibilidades que hemos discutido y que en realidad solo se basen en aspectos técnicos. Además de la estrategia de codificaciones teóricas para los casos faltantes (una estrategia que se acerca a la manipulación de datos), existen otras propuestas sobre cómo integrar el conocimiento teórico en la codificación de los casos faltantes o cómo reducir el número de los *simplifying assumptions* en la cual se basa la simulación para obtener un *resultado parsimonioso*. Estas propuestas incluyen la noción de las *hipótesis dirigidas* **(Ragin, 2008)**, la aplicación del *racionamiento contrafáctico* **(Ragin y Sonnett, 2004)** o la *organización del análisis causal en dos fases* **(Schneider y Wagemann 2007a)**.

Entonces ¿por qué, si estamos discutiendo la *diversidad empírica limitada* se presenta un problema, que es una cualidad del **QCA**? La cuestión es que el **QCA,** solo nos ayuda a explicitar el fenómeno y las dimensiones de la *diversidad empírica limitada* o diversidad limitada de una investigación. Este fenómeno, ciertamente se presenta en otras formas de investigación. Por ejemplo, en los estudios de caso, si tenemos cuatro casos, estos se compondrían al menos de cuatro características. Pero en realidad,

tendremos la necesidad de *16 configuraciones diferentes* y no solo de cuatro casos. De otra manera corremos el riesgo de que, o nuestras conclusiones se basen en *configuraciones no existentes*, o no sabemos identificar, cual de los cuatro factores sea en realidad la causa para lograr el *resultado*. Dicho lo anterior, se ve que no es tanto la *causalidad por conjunción* la problemática, sea en el método de Mill o en el *análisis comparado* en general, sino la *diversidad empírica limitada* o *diversidad limitada*, de la cual deriva nuestra incapacidad de resolver el nudo de la *causalidad por conjunción*.

Para la estadística, la *diversidad empírica limitada* o *diversidad limitada*, es igualmente un problema. O mejor dicho, la problemática con la estadística es que la *diversidad empírica limitada* o *diversidad limitada*, no es considerada un problema aunque exista en la propia estadística. Para afrontarla, los métodos estadísticos recurren a una estrategia ya comentada, la *parsimonia*. Por lo tanto, nuestros resultados en estadística se basan en una realidad empírica parcialmente inexistente y solo señalamos uno de los resultados posibles. Por ejemplo:

1. Suponga que deseamos entender dos de las causas de la inteligencia competitiva **I**, que son: la innovación del modelo de negocios del producto/servicio **P** y la detección de fuertes oportunidades de mercado, **S**.
2. Si el resultado de nuestra investigación, fuera el siguiente **Tabla 1.10**:

Tabla 1.10. Análisis de la diversidad limitada con métodos cuantitativos

Item	P	S	I
1	1	1	1
2	1	0	0
3	0	0	0

Fuente: Elaboración propia.

3. Con la ayuda de la estadística, se concluiría que existe una *correlación perfecta* entre la existencia de detección de fuertes oportunidades de

mercado **S** y causas de la inteligencia competitiva **I**, para dejar afuera la variable relativa a la innovación del modelo de negocios del producto/servicio **P**. Sin embargo, el problema con este resultado *perfecto* es que la tabla de arriba no está completa. En realidad la tabla debería conformarse de la siguiente manera. (Ver **Tabla 1.11**):

Tabla 1.11. Análisis de la diversidad limitada con métodos cuantitativos a tabla completa

Item	P	S	I
1	1	1	1
2	1	0	0
3	0	0	0
4	0	1	¿?

Fuente: Elaboración propia.

4. Las interrogantes sugieren, que nos hace falta información sobre qué sucedería con causas de la inteligencia competitiva **I** si existieran casos con un fuertes oportunidades de mercado **S** y falta de innovación del modelo de negocios del producto/servicio **P**. En efecto, nuestra conclusión sobre la correlación perfecta fuertes oportunidades de mercado **S** y las causas de la inteligencia competitiva **,** se mantendría solo si incluímos (**1**) allí donde están los signos de interrogación. Colocando (**0**), el resultado cambia y comprendemos que la falta de innovación del modelo de negocios del producto/servicio **P,** no está fuera, como se apreciará; los métodos estadísticos por el contrario, no incluyen la posibilidad de llevar a cabo este razonamiento sobre casos no existentes sino más bien, generan la solución más parsimoniosa.

Dado que la *solución parsimoniosa* es también una de las estrategias del **QCA**, lo anterior no es grave en sí mismo. El problema, es que las técnicas estadísticas, basadas en el software de uso frecuente, no muestran estas cir-

cunstancias. Así, se concluye que en general, dos puntos de vista del **QCA (Schneider y Wagemann, 2012)**:

1. *Que no es del todo una metodología nueva*. Utiliza un álgebra diferente al álgebra lineal, pero por lo demás es simplemente una versión más sistematizada de las estrategias tradicionales del *método comparado*. Nos ayuda a mostrar fenómenos como la *diversidad empírica limitada* o *diversidad limitada*. Otro ejemplo, es la existencia de *casos contradictorios*, es decir, si una y la misma *configuración* está relacionada *a casos con un resultado y a casos sin resultado* **(Ragin, 1987, p. 113)**. Las posibilidades de cómo afrontar una situación de este tipo son similares a aquellas para la *diversidad empírica limitada* o *diversidad limitada*; pero no se discutirá, dado que *el problema de los casos contradictorios*, solo existe en la versión *dicotómica* de la **csQCA** y no en la versión más prometedora, la **fsQCA**. Por lo tanto, el estudio del **QCA,** al ser una versión más sistematizada de las estrategias tradicionales del método comparado muestra ser además, importante para las investigaciones comparadas y que regularmente son inadvertidas por otras perspectivas metodológicas.
2. *Pero lo novedoso como metodología*, es su versión de los conjuntos difusos (*fuzzy-sets*). La aplicación de éstos nunca había sido considerada en las *metodologías comparadas* más tradicionales, simplemente porque su introducción en las matemáticas, es muy reciente **(Zadeh, 1965 y 1968)**. La *fuzzy álgebra*, es una formalización general de las reglas más específicas del álgebra Booleana **(Schneider y Wagemann, 2007a, p. 187; Klir et al., 1997; Ragin, 2000, p. 171.; Schneider y Wagemann, 2012)**. El **fsQCA** se basa en valores *fuzzy*, es decir difusos, y por lo tanto *no es una extensión del* **csQCA** *dicotómica*. El **csQCA** *dicotómico* por el contrario, es un caso especial del **fsQCA** más amplio. La novedad del **QCA** se encuentra precisamente en la versión de los *conjuntos difusos* (**fsQCA**) que permite aplicar una nueva álgebra haciendo *más refinados* los análisis comparados. Por ello, cada investigador que utiliza el *antiguo* **csQCA** se debe preguntar si es posible razonar con la ayuda de

las categorías difusas o sino es mejor recurrir a una comparación tradicional que no utilice el **QCA**. Esto no significa, que el **csQCA** sea una herramienta inútil ya que aún y cuando los conceptos son en realidad *dicotómicos*, los resultados de un **csQCA** se prefieren a la aplicación del **fsQCA**, porque el algoritmo del **csQCA** produce resultados que se interpretan de manera más simple. También, en el estudio del **QCA**, el **csQCA** debería preceder el análisis del **fsQCA**, porque los aspectos técnicos son mucho más claros en el **csQCA**, por lo que no es el estudio del **QCA** en general, sino de un caso especial.

QCA y el método comparado de Mill

Los fundamentos lógicos del *método comparado* que constituyen la base de **QCA**, pueden ser retrotraídos **(Berg-Schlosser et al., 2009, p. 2)** a los *cánones* de John Stuart Mill (1806-1873) filósofo, político y economista inglés, en su obra "*A system of Logic, Ratiocinactive and Inductive"*, especialmente a lo que se refiere al *método de la concordancia* y al *método de la diferencia* (que Mill combinaría en un *tercer método, el de la concordancia y diferencia*), del que se tiene:

1. *El método de la concordancia,* que afirma que si dos resultados iguales solo tienen una circunstancia común (divergiendo todas las demás circunstancias entre uno y otro), esa circunstancia común debe ser la causa que los provoca **(Mill, 1967, 1843, p. 450).**
2. *El método de la diferencia*, que afirma el sentido contrario, que si observamos dos fenómenos distintos cuyas circunstancias son iguales en todo salvo por una cosa, esta circunstancia divergente debe ser la causa que los explica **(Mill, 1967, 1843, p. 454)**.

Aunque ambos cánones reposan en supuestos rígidamente positivistas **(Berg-Schlosser, 2009, pp. 2-3)** sobre las relaciones causa-efecto, ambas constituyeron en su momento un valioso avance al mmento de eliminar factores *irrelevantes* en la explicación de las causación de los fenómenos.

Lo que el Dr. Ragin ha logrado crear **(Coverdill y Finlay, 1995, p. 458)**, es el desarrollo y refinamiento creados por **Mill (1967,1843)** a través de software **fsQCA**, siendo este determinista y que sigue procedimientos muy similares, para eliminar las *condiciones irrelevantes* siendo muy superior su desempeño a los métodos de Mill. De hecho, para juicio de los especialistas como **Schneider y Wagemann (2007a, pp. 4-5)**, los métodos de Mill no son capaces por sí mismos de ofrecer resultados satisfactorios, además de que su uso combinado, era escasamente realista con configuraciones empíricas más complejas a lo que ofrece la era informática, circunstancia superada en el **QCA** actual. Lo anterior básicamente por **(Mello, 2014, p. 59)**:

1. La incorporación de *tablas de verdad,* en donde se identifican todas las combinaciones posibles de condiciones, que permite identificar y abordar la diversidad limitada de los datos empíricos.
2. La introducción de la variante **fsQCA**, *que permite atribuir valores que oscilan entre* **(0)** *y* **(1)** y no solamente valores dicotómicos, así como de medidas separadas de ajuste *como la cobertura y la consistencia*, que permiten al **QCA** tomar en consideración la diferenciación en los efectos estimados y aportar *medidas de robustez*, en los análisis.

QCA y el Álgebra Booleana

Al **QCA** a menudo se le considera el paso intermedio entre lo cuantitativo y lo cualitativo. Aunque representa mucho más que eso, en lo cuantitativo se puede decir que emplea condiciones que varían la codificación numérica de las mismas, así como un tratamiento matemático, aunque no de tipo estadístico, sino basado en herramientas como la teoría de conjuntos o el álgebra Booleana.

George Boole (1815-1864) matemático británico, diseña un nuevo tipo de álgebra muy popular en el diseño de computadoras en el siglo XX y que ha logrado una gran aceptación en las ciencias sociales a partir del trabajo del Dr. **Ragin (1987)** adaptándolo como uno de los pilares que sostienen a **QCA** por:

“...el uso binario de datos, el uso de tablas de verdad para representar los datos, la adición y multiplicación Booleana, la lógica combinatoria, la minimización Booleana, la implicación y el uso de implicaciones principales, el uso de la Ley De Morgan para obtener soluciones, los conceptos de necesidad y suficiencia, así como la factorización de expresiones Booleanas...”

De acuerdo con **Ragin (2000, pp. 120-121)**, la lógica Boolena permite liberar al investigador del álgebra lineal que subyace a la metodología estadística y también de las restrictivas suposiciones de las investigaciones basadas en variables, proporcionando herramientas alternativas para *entender los casos como configuraciones que pertenecen a conjuntos.* Es más, incluso en los casos de dicotomizar los fenómenos, la lógica de minimización Booleana no desaparece. El enfoque Booleano (**Ragin, 1987, p. 101**), lo hace un instrumento ideal para comprender los patrones de *causalidad por conjuntura múltiple*, es decir, para comprender cómo distintas condiciones se combinan para producir un resultado y cómo un mismo resultado puede estar causando diferentes combinaciones de condiciones. Al mismo tiempo, el álgebra Booleana, aunque comienza con un sesgo hacia la complejidad, permite luego reconducir los análisis hacia la obtención de soluciones *parsimoniosas.*

Funciones de las técnicas QCA

De acuerdo con **Berg-Scholsser et al., (2009, p. 15)**, son cinco:

1. *Resumen de la información.* Al identificar el investigador las principales dimensiones, factores, características o condiciones de los casos, **QCA** es capaz de ofrecer sistemáticamente, las diversas combinaciones que se derivan, así como indicar en qué casos encajan dichas configuraciones, basados inicialmentente en las *tablas de verdad,* muy útiles al momento de construir taxonomías, tipologías y clasificaciones.

2. *Prueba en la consistencia de los datos*, cobrando una especial importancia el concepto de *contradicción*, la cual es la referencia a configuraciones de condiciones idénticas cuyos casos tienen *resultados contradictorios* **(Berg-Scholsser, et al., 2009, p. 15)**. Vea la **Tabla 1.12**

Tabla 1.12. Ejemplo hipotético de configuración contradictoria

Empresa clúster	Inversión I+D	Peso del sector industrial	Resultado de interés en innovación
	A	B	C
A	Alta	Alto	Alto
B	Baja	Alto	Bajo
C	Alta	Alto	Bajo

Fuente: Elaboración propia.

donde se ven las empresas del clúster **A** y **C** que comparten una misma configuración de condiciones: alta inversión en I+D y un sector industrial de peso alto. Sin embargo, mientras que **A** logra alcanzar un alto desarrollo económico, no sucede igual con **C.** Nos enfrentamos de tal manera a una contradicción lógica, puesto que condiciones idénticas conducen a resultados opuestos, apuntando así a una deficiencia en el marco teórico o en el análisis que la investigación debe subsanar. **QCA** permite identificar estas contradicciones y abordarlas de manera efectiva.

3. *La prueba de hipótesis o teorías existentes*. Es probablemente, una de las funciones fundamentales del **QCA** y posiblemente, la más usada. Como ocurre con otras técnicas de investigación social, **QCA** permite falsear hipótesis, sometiendo sus enunciados al análisis de la evidencia empírica. En este sentido, será preciso tener en cuenta la naturaleza configuracional del análisis **QCA**, pues afectará a la manera en que las hipótesis deben ser formuladas **(Berg-Scholsser et al., 2009, pp. 15-16)**.

4. *Prueba rápida de supuestos*. **QCA**, permite realizarlo sin necesidad de que vengan precedidas por una teoría completa, mediante la especifica-

ción de una fórmula que las refleje y que será analizada a la luz de una *tabla de verdad* (*Ibídem*).

5. *Permitir el desarrollo de nuevos argumentos teóricos*. Los análisis **QCA** producen la identificación de una serie de patrones causales, es decir, *configuraciones* de condiciones que causan un resultado. Estos patrones causales, pueden ser considerados como narrativas acerca de la causación de los resultados y en este sentido, pueden integrarse posteriormente al *corpus* del conocimiento del campo en forma de aportaciones teóricas (*Ibídem*).

El enfoque configuracional y cómo formular hipótesis

En los estudios que utilizan técnicas estadísticas, las hipótesis se formulan habitualmente en una forma que refleja la naturaleza probabilística y correla-cional de los resultados que se esperan obtener: *si* **X** *aumenta / disminuye,* **Y** *tenderá a aumentar / disminuir.* Sin embargo, en **QCA** las relaciones causales son deterministas y basadas en las ideas de *necesidad y su iciencia* del tipo **(X)** *es condición necesaria / su iciente para* **Y** **(Roberts Clark et al., 2006)**. En este sentido, se usa el concepto de *hipótesis asimétricas,* que puede resultar útil porque permite aunar las ideas de *asímetría y necesidad/ su iciencia* con las que opera **QCA** y entender la vinculación conceptual entre ambas. Se debe entender que **(Roberts Clark et al., 2006, p. 315)**:

a. La *simetría* se refiere a causas que son *necesarias y su icientes,* por ejemplo, cuando **X** aumenta se produce un aumento de **Y** y cuando **X** disminuye se produce una disminución de **Y.**

b. La *asimetría*, por el contrario, se refiere a causas que son, o bien *necesarias* para una consecuencia, por ejemplo, cuando **X**)está presente, **Y** puede ocurrir o no, pero si **X** está ausente **Y** no ocurrirá; o bien, *su icientes,* o sea, cuando **X** está presente **Y** ocurrirá, pero cuando **X** está ausente **Y** puede o no ocurrir.

c. Las hipótesis en **QCA** deben ser *configuracionales* **(Schneider y Wagemann, 2010, p. 4)**, es decir, plantear hipotéticamente la causa-

lidad como resultado de la interacción entre condiciones, esto es, por ejemplo, hipotetizar que la presencia conjunta de **A** y **B** es *su iciente* para la ocurrencia de **Y.** La naturaleza configuracional de **QCA** se debe tomar en cuenta al momento de analizar los resultados. Suponga del caso anterior que cumple; no bastaría decir que la condición **A** tiene impacto sobre **Y** y que lo mismo ocurre de forma separada, con la condición **B.** La solución del análisis está diciendo que ambas condi-ciones, para tener un impacto causal, debían darse conjuntamente. ¿Qué nos está contando dicho resultado sobre nuestros casos? ¿Por qué no era suficiente que se dieran **A** o **B** por separado para producir **Y**? ¿En qué consiste esta interacción y por qué es tan importante desde una pers-pectiva teórica? El uso de una técnica configuracional exige, de análisis e interpretaciones *con iguracionales* que, por lo general, pueden ser entendidas como *historias o narrativas* que se ocultaban detrás de los casos y que hemos sido capaces de desentrañar como *patrones causales* que conducen a resultados, en los que diferentes fenómenos interactúan entre sí produciendo el fenómeno analizado.

Razones por las que QCA representa un alcance ambiguo

Tratar el aspecto cualitativo del **QCA,** no es tan fácil como pudiera pensarse. En la tradición francesa, al **QCA** se le denomina "*Analyse Quali-Quantitative Comparée*" (**AQQC**) **(De Meur et al.,2002)**, integrando explícitamente el adjetivo *cuantitativo* en la abreviación de la técnica. **De Meur (et al., 2002)** ubican el **QCA** entre los métodos cualitativos y cuantitativos, **Ragin (1987)** lo ubica más allá de esta distinción (*moving beyond qualitative and quantitative strategies*) de hecho, no se trata de discutir que cuán útil es la distinción entre perspectivas cuantitativas y cualitativas o entre investigación orientada a casos (*case-oriented*) y la orientada a variable (*variable-oriented*) **(Ragin 1987; 2004; Ragin y Zaret,1983; Bartolini, 1993, p. 173, nota 9)**.

Es de observar que **Ragin (1987)**, inicialmente evita las siglas *la expresión Qualitative Comparative Analysis* o la abreviación (**QCA**) no se

encuentra en dicha obra. Se encuentra, solo posteriormente con frecuencia, aunque es fácil establecer dónde y cuándo fue introducida. Más tarde, Ragin utiliza la expresión *case-oriented* **(Ragin, 2000, p. 23)** o *diversity-oriented* **(Ragin, 2000, p. 19)**, que para una técnica como el **QCA** es quizá, más acertado. No obstante, tales variaciones en lo que a terminología se refiere, Ragin declara explícitamente que, refiriéndose a la discusión (*estadounidense*), entre estudiosos cuantitativos y cualitativos, el **QCA** *es parte de las perspectivas cualitativas* **(Ragin, 2000, p. 13)**. Aún así, todavía queda la duda de si la terminología en francés, que habla de un análisis *cuali-cuantitativo* no sea la más correcta, dado el uso de dos tipos de álgebra, *la aplicación de parámetros de evaluación de la calidad del análisis y la codificación de conceptos con cifras*. Sobre todo en el caso del **fsQCA**, la terminología es utilizada de una manera en la cual los resultados del análisis se asemejan mucho al estilo de los métodos utilizados en estadística **(Schneider y Grofman, 2006)**. Resulta paradójico, que el libro en lengua francesa que introduce la expresión de análisis *cuali-cuantitativo* **(De Meur et al., 2002)** trate solo el **csQCA** y no discuta el **fsQCA.**

Razones por las que QCA es aún una técnica cualitativa

Existen siete razones (quizás aún discutibles) que justifican el **QCA** como una de técnica *cualitativa*, a saber **(Wagemann, 2012)**:

1. En primer lugar, el **QCA** se concentra sobre todo en las características de cada uno de los casos de manera individual. Varios pasos del proceso técnico, interrelacionan los resultados obtenidos con los casos en los cuales se basa el análisis **(Schneider y Wagemann, 2007b).** La última forma de control, recomendados, debe referirse de nuevo a la explicación de cada uno de los casos de manera individual.
2. En segundo lugar, los casos son *des-estructurados* en configuraciones en relación a sus características. Este procedimiento, comparte la perspectiva holística con los estudios de caso. Las características de cada

uno de los casos, no deben desaparecer detrás de las variables sino que deben permanecer en el centro de la investigación.

3. En tercer lugar, las codificaciones no se guían a través de *reglas cuantitativas*. Aunque paradójico al inicio, la única manera para decidir sobre las dicotomías en **csQCA** o sobre valores difusos (*fuzzy*) en la **fsQCA** es recurriendo a la teoría **(Hall, 2003, p. 389; Mahoney, 2003, p. 347; Ragin, 2000, p. 150)**. Tal necesidad presupone un conocimiento profundo *(cualitativo*) de cada caso y de la materia estudiada. Esto es, el investigador decide el grado de pertenencia, o no, a un concepto.
4. En cuarto lugar (y con relación al punto anterior), las álgebras de la **QCA** (el álgebra Booleana y la *fuzzy-algebra*) no son álgebras que cuenten los objetos bajo análisis. Por el contrario, se observa primero si una cualidad existe o no (y en el extremo, cuanto existe). De esta forma, el resultado de un **QCA** no cambia si una configuración existe una vez o más de una vez, es decir, *si existe un solo caso para cierta configuración o si la configuración representa más de uno solo caso*.
5. Es característico y por lo tanto recomendable de la **QCA,** que la redefinición de los casos (*y no solo de una muestra sino de todo el universo de casos*), de las condiciones, e incluso el resultado, sean parte del proceso analítico. En el caso de una *re-especificación del universo de casos* (o población) se trata de dar una nueva respuesta a la famosa pregunta *¿en qué cosa el caso es un caso?* **(Collier y Mahoney, 1996, pp. 4 y 38; Ragin, 2000, p. 53)**. Esto no significa que se deban manipular los datos. Por el contrario, se trata de un proceso de aprendizaje a partir de la evidencia y de la reformulación de ideas nuevas y de inventar nuevas partes de la investigación con la ayuda de la evidencia **(Munck, 2001, p. 119; Ragin, 1994a, p. 76; 2004, p. 126)**. En este sentido la **QCA** está también en consonancia con los estudios de caso, como una estrategia de investigación que permite adecuar el diseño de la misma. En estadística por el contrario, la revisión de un modelo deriva sobre todo de las carencias y/o fallas matemáticas de otro. De allí que *la re-especificación del universo de casos, nunca es aconsejable en los métodos cuantitativos.*

6. La perspectiva específica sobre la causalidad, está muy cercana a la idea de causalidad contenida en la investigación cualitativa en general. Sobre todo la *centralidad de la complejidad causal* se encuentra en diversos estudios de caso e incluso en investigaciones muchos más cualitativas (no cuantitativas o estadísticas).
7. **QCA** en el caso ideal, trata de generar *conclusiones determinísticas*. Las medidas de *consistencia* y de *cobertura* nos indican desviaciones del determinismo ideal. Sin embargo, no se trata de afirmaciones sobre la probabilidad de que el resultado se presente o no. Es en este aspecto donde reside la diferencia con los métodos estadísticos, (que son precisamente la aplicación de cálculos sobre probabilidades). Por ello la literatura sobre la **QCA** evita llamarlos resultados *no-determinísticos probabilísticos* (que tienen una *consistencia* inferior a **1**). Por el contrario, utiliza el término *posibilístico*.

Como se puede observar, no obstante siendo en extremo formalizado y basada sobre el álgebra, la **QCA** comparte muchos de los presupuestos con la investigación que estamos habituados a denominar *cualitativa*, que con los métodos estadísticos, por lo que está bien identificarlo aún, como una técnica cualitativa.

Técnicas dominantes QCA

Existen hoy en día, tres técnicas en **QCA**, que conviven con algunas otras relativamente novedosas y por ello aún minoritarias, pero no de menor potencial; tal es el caso de **TQCA** (*Temporal* **QCA**) o **gsQCA** (*generalised* **QCA**) **(Medina et al., 2017)**. Ver **Tabla 1.13.**

Tabla 1.13. Técnicas y tipo de condiciones en QCA

Técnica	Condiciones	Software	Link
Crisp-set QCA	Dicotómica	fsQCA 3.0	http://www.socsci.uci.edu/~cragin/fsQCA/software.shtml
		Package QCA	https://cran.r-project.org/web/packages/QCA/index.html
		STATA (Package Fuzzy)	https://download.stata.com/download/
Multi-Value QCA	Multicotómicas (aunque también se admiten Dicotómicas, especialmente para el resultado)	TOSMANA v 1.52	https://www.tosmana.net/
		Package QCA	https://cran.r-project.org/web/packages/QCA/index.html
Fuzzy-Set	Difusas (aunque también se admiten Dicotómicas, excepto para el resultado)	fsQCA 3.0	http://www.socsci.uci.edu/~cragin/fsQCA/software.shtml
		Package QCA	https://cran.r-project.org/web/packages/QCA/index.html
		STATA (Package Fuzzy)	https://download.stata.com/download/

Fuente: Medina et al., (2017)

Como se podrá apreciar, las diferencias entre sí, se basan sobre todo por el tipo de condiciones que incluyen el análisis (*dicotómicas*, *multicotómicas y difusas respectivamente*), pero a todas las une una serie de características comunes, tales como su naturaleza *configuracional, su lógica de minimización Booleana* y su relación con la *teoría de conjuntos*. Las técnicas **QCA** no utilizan un lenguaje centrado en variables dependientes e independiendientes, sino reespectivamente, en resultados y condiciones, de modo que se entiende que es precisamente la interacción entre dichas condiciones (conocidas como configuraciones) y no su impacto aislado, la que provoca el resultado de interés. Estas características, diferencian a **QCA** de otras

técnicas y enfoques en las ciencias sociales, particularmente de las técnicas probabilísticas, permitiendo un abordaje de la investigación que parte de supuestos espistemológicos distintivos muy útiles.

El futuro de QCA para mejorar

Como se ha visto hasta el momento **QCA** se caracteriza por (**Wagemann, 2012**)**:**

1. Logra sostener hipótesis que no se basan en correlaciones sino *bajo las premisas de condiciones suficientes y necesarias*, y sobre las condiciones **INUS** (*lnsufficient but Necessary part of a condition which is itself Unnecessary but Sufficient for the result*) y las condiciones **SUIN** (*Sufficient, but Unnecessary part of a factor that is Insufficient, but Necessary for the result*).
2. Se ha concluído, que la **QCA** logra también analizar la *complejidad causal* inherente a muchos de los estudios comparados.
3. La novedad de **QCA**, consiste en una formulación sistemática de los principios fundamentales de la investigación comparada en general: en una *fuerte atención a los problemas de la investigación comparada*, como lo ilustra el ejemplo de la *diversidad empírica limitada o diversidad limitada*; y en la introducción de una *nueva álgebra*, la álgebra difusa (*fuzzy algebra*) en la metodología comparada.
4. Finalmente, se ha logrado la justificación de llamar a la **QCA** una metodología cualitativa.

Sin duda, la **QCA** es una metodología joven, aunque después de la revisión del algoritmo del 2003, de acuerdo a **Ragin (2004)**, por aportar más evidencias de que la técnica pueda ser considerada estable e irrevocable. Por lo tanto, podemos suponer que todavía existan muchos aspectos por mejorar. Brevemente trataremos de enlistar los pasos más urgentes a considerar **(Schneider y Wagemann, 2012)**:

1. La idea de crear una versión que logre tener en cuenta los conceptos multi-lineares se inició con la llamada **mvQCA (Cronqvist, 2005)** debe desarrollarse todavía. Sobre todo debe elaborarse de manera que sea aplicable a la realidad de la investigación empírica y debe fundamentarse mucho mejor en la teoría de conjuntos. Si las diversas variantes de la **QCA** no están interconectadas con una base común de la teoría de conjuntos se corre el riesgo de desligar el nexo entre las mismas;
2. La **QCA** aún es incapaz de modelar procesos temporales los cuales son, como sabemos, importantísimos en la *investigación comparada* **(Bartolini, 1993; Pierson, 2003)**. El desarrollo en este sentido, es deseable y aún incipiente **(Caren y Panofsky, 2005; Ragin y Strand, 2008)**. El efecto del momento, puede variar dependiendo del contexto. Las cuestiones relativas a la dependencia del tiempo y del contexto podrían modelarse muy bien con un método sensible a los *factores por conjunción* y a la *causalidad compleja* como la **QCA**;
3. También, se deben producir mucho más *líneas-guía* sobre cómo producir los valores difusos (*fuzzy*) para cada caso. De otra manera, se corre el riesgo de la ambigüedad y *todo se valdría*, incluso la manipulación de los resultados. Pasos iniciales en este sentido ya se han llevado a cabo **(Ragin, 2007; 2008)**;
4. La discusión sobre cómo enfrentar el fenómeno de la *diversidad empírica limitada* o *diversidad limitada*, debe continuarse, no solo considerando la **QCA**, sino también otros métodos. En efecto, es un problema aún no resuelto porque los casos que no existen *simplemente no existen* y no se los deben inventar; por ello se deben buscar estrategias sobre cómo reducir el riesgo para no llegar a conclusiones equivocadas debido a la falta de casos;
5. El último punto es quizá el más importante, porque la comunidad de los usuarios de la **QCA** está creciendo constantemente. Esto es positivo porque ayuda a difundir el método. Pero el aspecto negativo es que la **QCA** es a veces aplicado de manera superficial, automática e incluso de manera incorrecta técnicamente hablando. La otra cara de la moneda del uso frecuente es que el análisis desarrollado de manera equivocada

es una publicidad negativa para el método mismo. De allí la necesidad de un código explícito, en el cual se resuman los estándares para una **QCA** de alta calidad y confiabilidad. De esta manera los lectores, pero también los editores, dictaminadores y sinodales de trabajos de investigación, así como los estudiantes tendrían un instrumento para juzgar no solo sobre la calidad del análisis, sino también sobre la validez de los resultados obtenidos. Hace algunos años se trató **(Ragin y Rihoux, 2004, p. 6.; Yamasaki, 2003, p. 3; Schneider y Wagemann, 2007b; 2010)**; pero la idea de que la fuerza real de un instrumento metodológico depende también (y, sobre todo) de su correcta aplicación, no está del todo bien difundida. Podemos por ahora concluir que la **QCA** es prometedora para los *análisis comparados* y que sería recomendable una difusión más amplia a nivel internacional. Evidentemente se trata de un trabajo en proceso (*work in progress*) con algunos problemas todavía por resolver. Pero el método es aplicable (y de hecho se aplica) y nos puede ayudar a descifrar de mejor forma a la sociedad y su desenvolvimiento en las prácticas de la innovación.

Hacia el fsQCA en las ciencias sociales y más allá

El análisis cualitativo comparativo de *datos difusos* (**fsQCA**. *Fuzzy Set Qualitative Comparative Analysis*), desarrollado por el científico social Charles C. Ragin, es una metodología para obtener interpretaciones de datos en argumentos de *condiciones causales*, que están asociados con los casos. El **fsQCA** busca establecer conexiones lógicas entre las combinaciones de *condiciones causales* y un resultado deseado, el resultado son reglas (*configuraciones tipológicas*) que resumen (*summarize*) o mejor dicho, describen, la suficiencia entre subconjuntos de todas las combinaciones posibles, de las *condiciones causales* (*o sus complementos*) y el resultado. No es una metodología derivada de las Matemáticas, aunque las utiliza, se encuentra escasamente descrita. Estas causas y sus complementos, se convierten en los antecedentes en las combinaciones causales **fsQCA** y, en cada combi-

nación causal, siempre están conectados por la operación lógica **AND** que se modela utilizando el operación mínima.

Un estudio previo de **Vis (2012)** de **53** muestras, realiza un comparativo entre los resultados de la regresión lineal y **fsQCA** donde demuestra, que si bien cada enfoque tiene méritos y deméritos, **fsQCA** conduce a una comprensión más completa de las condiciones bajo las cuales ocurre el resultado. A diferencia de otros métodos cuantitativos, que se basan en la correlación, **fsQCA** busca establecer conexiones lógicas entre combinaciones de *condiciones causales* (*causalidad por conjunción*) generalmente conectadas por la palabra lógica **AND**, asociadas a un resultado; éste conforma un conjunto reglas que resumen el ser suficiente entre subconjuntos de todas las combinaciones posibles de las condiciones de causalidad (*o sus complementos*) y el resultado. Las reglas, están conectadas por la palabra lógica **OR** a la salida.

Cada regla es una posible trayectoria causal o solución, que parte desde las *condiciones causales* iniciales, hasta el resultado y representa la causalidad equifinal, es decir, diferente combinaciones causales que conducen al mismo resultado. Es así, que ya existen intentos por clarificar los pasos de aplicación de **fsQCA**, como los mencionados por **Mendel y Korjani (2010) y Korjani y Mendel (2012)**. Ver **Tabla 1.14.**

Tabla 1.14. Los 13 Pasos sugeridos para aplicar fsQCA

Descripción
Paso 1. Seleccione el resultado deseado (*outcome*) y sus casos apropiados Elegir el resultado deseado, se basa en la aplicación; sin embargo, identificar los casos apropiados para el resultado deseado, es un desafío. Para elegir los casos apropiados, se necesita saber cuál es la definición de qué es un caso. Según **Ragin (2009)**, la pregunta ¿cuál es el caso? puede tener diferentes respuestas. De hecho, el primer paso, en las investigaciones orientadas a casos, es identificar las mejores posibles instancias del fenómeno a explicar y luego estudie estos casos en gran profundidad. Lo anterior, es a fin de obtener diferentes opciones de casos para diferentes tipos de estudios, por ejemplo: **a)** Los que van desde un estudio para el cual hay solo un caso, **b)** Los que van a un estudio en el cual, hay un conjunto de casos para el mismo resultado, **c)** Los que van a un estudio para el cual hay casos negativos y positivos para el mismo resultado, **d)** Los que van a un estudio que utiliza a toda la población (tal estudio, busca generalizaciones sobre la población). Tenga en cuenta que un caso positivo, es uno para el cual el resultado deseado está fuertemente presente y un caso negativo es aquel, para el cual el resultado deseado puede no estar presente en absoluto o está débilmente presente. La elección de los casos apropiados debe hacerse primero, y esta elección no tiene que hacerse desde el inicio, es decir, puede ser modificado durante todo el procedimiento **fsQCA**.
Paso 2. Selección *k* de *condiciones causales* Si una condición es descrita por más de un término, trate cada término como una *condición causal* independiente. En esta etapa, un subconjunto de todas las posibles *condiciones causales*, es seleccionada.
Paso 4. Evalúe la función de pertenencia (MF. Membership Function) de las muestras apropiadas de caso N Los casos **MF**, pueden generarse a partir del algoritmo Fuzzy-C-Means (**FCM**). Si aplicamos **FCM** a cada variable o resultado deseado por separado **(Bezdeck, 1981)**. Se recomienda realizar primero el paso 4 antes del 3, debido a cómo obtener los **MF** para las *condiciones causales* y la salida deseada.
Paso 3. Trate la salida deseada y las *condiciones causales* como *datos difusos* (*fuzzy-sets*) y determine las MF para ellos Se regresará al paso 4 para reevaluar los **MFs** de todos los casos apropiados **N.**

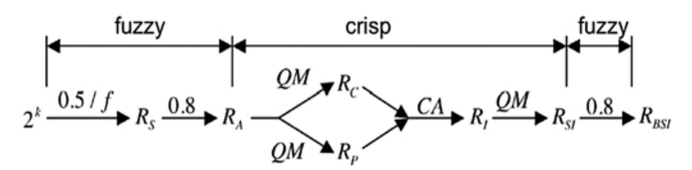
Paso 5. Crear 2^k combinaciones causales candidatos Estas son las reglas e interprete cada una como una esquina espacial de $\mathbf{2}^k$ vectores del espacio dimensional.
Paso 6. Calcule el MF de cada uno de los 2^k candidatos de combinaciones causales, en todos los casos apropiados Mantenga solo las combinaciones adecuadas de los **(Rs)** (*firing levels surviving rules*), cuyos valores **MF** son mayores a **0.5** para un número de casos, es decir, mantenga adecuadamente representadas las combinaciones causales, que están más cerca de las esquinas (no al contrario). Establezca umbrales de frecuencia adecuados.
Paso 7. Cálculo de *consistencias* (*subsethoods*) de las combinaciones causales sobrevivientes y mantenga solo aquellas cuyas *consistencias* sean mayores a 0.8
Paso 8. Utilice el algoritmo Quine-McCluskey (QM) Esto es para obtener soluciones complejas (implicaciones principales) y soluciones *parsimoniosas* (implicaciones principales mínimos).
Paso 9. Análisis contrafactual Usar conocimiento empírico sustantivo adicional, obtenido por parte del investigador experto, para realizar un análisis contrafactual en cada término de las soluciones complejas (una a la vez), pero limitado por cada término de las soluciones *parsimoniosas* (una a la vez), para obtener soluciones intermedias.
Paso 10. Desempeño del algoritmo (QM) Para obtener soluciones intermedias simplificadas.
Paso 11. Retención de soluciones intermedias Retenga solo aquellas soluciones intermedias simplificados, cuyas *consistencias* son aproximadamente mayores que 0.80 para considerarlas como, soluciones intermedias simplificadas creíbles.
Paso 12. Conecte las soluciones simplificadas intermedias posibles Relacionarlas con las mejores instancias de cada solución, es imprescindible.
Paso 13. Calcule la *cobertura bruta* (*raw coverage*) Haga el cálculo para cada solución. Esto representa el porcentaje de casos que soportan cada solución intermedia simplificada posible.

Fuente: Mendel y Korjani (2010); Korjani y Mendel (2012).

Dados los varios pasos que se ejecutan en **fsQCA** donde las combinaciones causales originales $\mathbf{2^k}$ se reducen a $\boldsymbol{R_{BSI}}$ (*reduced believable simplified intermediate summarizations*), **Mendel y Korjani (2010)**,

ofrecen un mnemónico para su mejor interpretación en la **Figura 1.5,** donde el énfasis se encuentra en las etapas difusa y nítida.

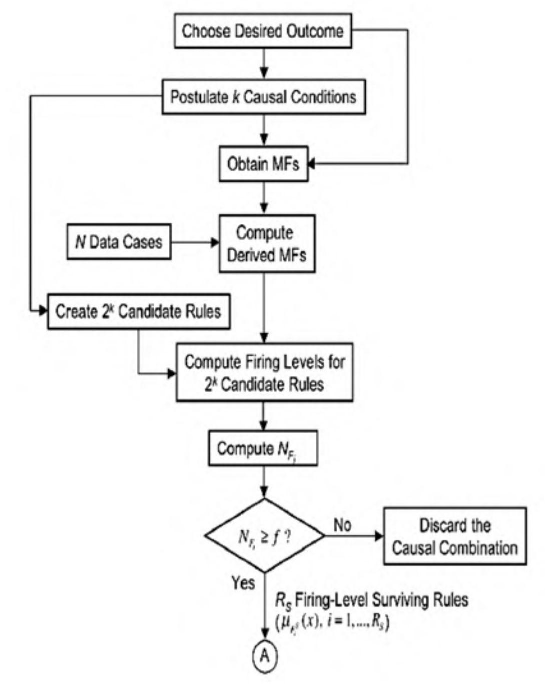
Figura 1.5. Mnemónico de los procesos fsQCA



Fuente: Mendel y Korjani (2010).

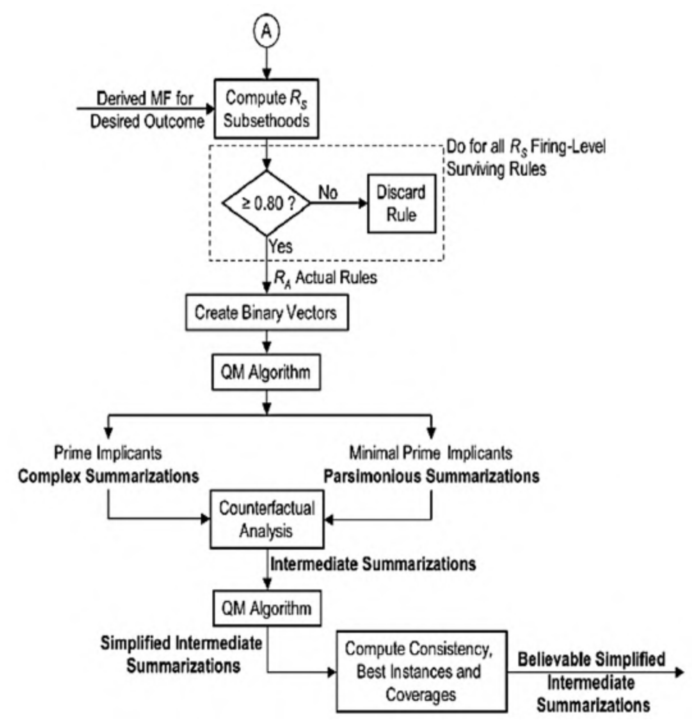
Además, se encuentra una aportación al describir las diferentes etapas en un diagrama de flujo que corresponde a los pasos mencionados en la **Tabla 1.14**. Ver **Figura 1.6** y **1.7**

Figura 1.6. Diagrama de flujo Pasos 1 al 6 de 13 para aplicar fsQCA



Fuente: Mendel y Korjani (2010); Korjani y Mendel (2012).

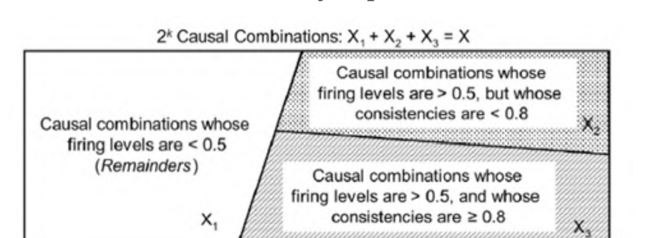
Figura 1.7. Diagrama de flujo Pasos 7 al 13 de 13 para aplicar fsQCA



Fuente: Mendel y Korjani (2010); Korjani y Mendel (2012).

Así también, como se deben visualizar las $\mathbf{2^K}$ combinaciones causales cuando se particionan, en tres conjuntos que no se sobreponen. Ver **Figura 1.8.**

Figura 1.8. 2^K Combinaciones causales y su particionamiento en tres conjuntos



Fuente: Mendel y Korjani (2010).

De hecho, **Mendel y Korjani (2010)** describen la posibilidad de que **fsQCA** realice resumenes linguísticos (*linguistic summarization*) ya que, de acuerdo a su punto de vista, el creador original, el Profr. Charles Ragin solo considera que **fsQCA** describe lo que está sucediendo en el transcurso de una colección de *condiciones causales* y el resultado. Es en la obra, por cierto de **Rihoux y Ragin (2009, p. 15)** en su figura 14 en la que se reconoce al resumen datos (*summarizing data*), como uno de los cinco tipos de usos, de las técnicas de QCA. En consecuencia, ahora parece legítimo utilizar **fsQCA**, como base para lograr el resumen lingüístico (*linguistic summarization*). Los otros cuatro usos para **QCA**, son:

1. Verificación de la coherencia de datos,
2. Verificación de hipótesis de teorías existentes,
3. Prueba rápida de conjeturas y
4. Desarrollo de nuevos argumentos teóricos.

De inicio, la técnica **fsQCA** utilizó simples nombres para identificar a las *condiciones causales*, tales como: empresa, industrial, electrónica, urbana, clúster, innovación, sin embargo, a medida que se ha requerido cada vez más su uso, se ha incrementado una serie de adjetivos, tales como: alta innovación, baja presencia de mercado, altos ingresos por innovación, etcétera. Si una *condición causal* tiene más de un término que lo describe (por ejemplo, alta innovación y altos ingresos por innovación), entonces cada término se trata como una *condición causal* independiente. Estas causas o sus complementos se convierten en antecedentes en el combinaciones causales **fsQCA** y, en cada combinación causal, siempre están conectadas por el operador lógico **AND**, que se modela utilizando la función **min** (mínima). Para **Ragin (2008)** el **fsQCA,** se usa separadamente de cada una de las salidas posibles, así como sus complementos. Por ejemplo, en un estudio de empresas generadoras de innovación hay dos tipos de estudios (*configuraciones tipológicas*), las de innovación y las que no poseen innovación. Generalmente hay una asimetría causal *(causal assymetry*), entre

los dos **fsQCA,** lo que significa que generalmente, no son los complementos de las combinaciones causales asociadas con la supervivencia de sistemas de innovación que estén asociados, con el colapso de los mismos. ¿Cuántos términos de resultado **fsQCA** se deben realizar? La respuesta depende del investigador.

Inicialmente, se pensó que tendría que realizarse un **fsQCA** separado para cada término de una *condición causal*. Hacer esto, no solo habría llevado a una explosión en el número de causalidades generadas por **fsQCA**, sino que también, habría llevado a posibles resúmenes contradictorios (por ejemplo, uno para baja permeabilidad y otro para alta permeabilidad) que de alguna manera tendrían que ser resuelto. La experiencia, ha recomendado que cuando use una *condición causal* que tenga dos términos asociados (por ejemplo, bajos y altos ingresos) trate cada término como una *condición causal* separada. Aquí, se adopta dicho enfoque.

Es interesante observar, las distintas opciones potenciales de desarrollo del **fsQCA,** ya que en los trabajos de **Mendel y Korjani (2010)** lo explican de una manera muy cuantitativa, algo que no se encuentra en la literatura existente, y que se necesita para el uso de las áreas de ingeniería **(Marks et al., 2018)**.

Actualmente, el **csQCA** y el **fsQCA** constituyen versiones superiores del **QCA**. Obviamente parece que el **fsQCA** tiene mayores posibilidades de aplicación. Mientras tanto se está creando una tercera versión del **QCA**, llamada multi-value **QCA** (**mvQCA**) **(Cronqvist y Berg-Schlosser, 2008)** la cual considera conceptos multinominales (por ejemplo, el concepto de *familia partidista*, el cual no es un concepto dicotómico ni un concepto ordinal, sino que es difuso). Empero, el **mvQCA** tiene todavía límites. Por un lado, es difícil enmarcar la **mvQCA** en términos de una teoría. Por otro, tiene varios problemas técnicos (por ejemplo, en relación a la *diversidad empírica limitada*) que todavía no han sido resueltos **(Schneider y Wagemann, 2012; Vink y Van Vliet, 2009)**.

Críticas al fsQCA

Ciertamente el **fsQCA** ha recibido críticas, tales como:

- En relación con que los valores difusos, no se basan en una *teoría de medición,* a lo que se respondió que, una estandarización de las decisiones sobre los valores difusos, no es posible. Esto es, porque representan conceptos muy complicados de las ciencias sociales, siempre diferentes entre sí.
- Otras críticas, se han referido al mismo algoritmo, del cual existe una nueva versión desde 2003 **(Ragin, 2006)**. En particular, el nuevo algoritmo enfatiza la construcción de *tipos ideales* (**Morlino, 2005, p. 76**) a través de los valores difusos en el sentido de un espacio de atributos.
- Incluso, es posible *medir* la pertenencia de un caso a un tipo ideal a través de una simple operación de álgebra difusa.
- También, los parámetros de la consistencia (*consistency*) y de la *cobertura* (*coverage*) **(Ragin, 2006)** sustituyen los elementos de la versión anterior **(Ragin, 2000, pp. 109 y 223)** que eran incompatibles con la lógica del análisis (ahora, *fuzzy adjustments* y *benchmark proportions*). Sobre todo, la introducción de la consistencia resultó en una nueva versión del algoritmo **(Ragin y Pennings, 2005)**. Por lo que le queda un futuro prometedor de desarrollo.

CAPÍTULO 2.

Software QCA y configuración fsQCA

En esta capítulo, se abordará la descripción y evaluación del software más utilizado para realizar análisis **QCA** así como la configuración del software **fsQCA** versión **3.0** diseñado por el Dr. Ragin y el Ingeniero Drass **(Rihoux y De Meur, 2009)** desde 1987.

Evaluación del software QCA

Probablemente, en relación directa con el creciente interés y uso de las técnicas **QCA**, en los últimos años estamos asistiendo a una verdadera oleada de creación de nuevos sistemas de software, para realizar este tipo de análisis, así como de actualización del software tradicional. Aunque se realizará la configuración e interpretación de resultados del software **fsQCA** versión **3.0** como una derivación de los dedicados a las técnicas **QCA,** merece la pena destacar la utilidad general, a modo comparativo, entre los diferentes programas para que el investigador conozca el potencial de cada uno de ellos (véase la **Tabla 2.1**).

Tabla 2.1. Resumen de las características de los distintos software para ejecutar QCA

Item	Características analíticas	fsQCA 3.0	TOSMANA 1.52	STATA	R
Esencial					
1	Tabla de verdad codificable	**X**	—	**(X)**	—
2	Solución intermedia	**X**	—	—	**X**
3	Exclusión automática de supuestos insostenibles	—	—	—	—
4	Parámetros de ajuste	X	—	**X**	**X**

5	Análisis de suficiencia en **csQCA**	**X**	**X**	**X**	**X**
6	Análisis de suficiencia en **fsQCA**	**X**	**(*)**	**X**	**X**
7	Ejecución de **mvQCA**	—	X	—	X
8	Análisis de necesidad	**X**	—	**X**	**X**
Visualización de resultados					
9	Diagrama de Venn	—	**X**	—	**X**
10	XY Plot	X	X	X	X
11	Identificación de casos en soluciones	**X**	**X**	**X**	**X**
12	Listado de casos en las filas de la tabla de verdad	**X**	**X**	—	**X**
13	Listado de supuestos simplificadores	—	X	—	X
Atributos					
14	Calibración difusa (*fuzzy-set*)	**X**	—	**X**	**X**
16	Basado en la sintaxis	—	—	**X**	**X**
16	Gratuito	**X**	**X**	—	**X**
17	Valculador Booleano	—	**X**	—	**(X)**
18	Plataforma abierta a desarrolladores	—	—	**X**	**X**
19	PRI y PRODUCT	**X**	—	—	**X**
20	Medidas estadísticas de ajuste	—	—	**X**	—
21	Compatible con todos los sistemas operativos	(+)	X	X	X

Nota:
(X) Significa que la opción prevista en paquetes específicos, pero es posible desempeñarla dentro del entorno del software.
(*) Significa que el software permite apoyo básico a la realización de análisis difuso (*fuzzy*).
(+) Windows y Mac.

Fuente: Schneider y Wagemann (2012, p. 283) con adaptación propia.

Para ello, seguimos en esta sección fundamentalmente los trabajos de **Schneider y Wagemann (2012, p. 282)** aunque con información actualizada para cubir las últimas innovaciones técnicas. Es importante destacar que, como los autores dicen, ningún software es el mejor para realizar análisis de **QCA**, aunque la ventaja de conocer las fortalezas y debilidades de cada uno de ellos estriba en su combinación **(Schneider y Wagemann,**

2012, p. 284). Cabe decir, que se trata de programas en constante evolución y desarrollo; su descarga y actualización puede encontrarse en el sitio web www.compasss.org.

El software más común para *los análisis configuracionales* ha sido tradicionalmente **fsQCA 3.0**. A juicio de **Rihoux et al., (2009a, p. 176),** este programa consiguió superar los problemas de **QCA-DOS 2.0 y 3.0**, operados durante las últimas décadas del siglo pasado en entorno **DOS**. El software permite ejecutar análisis **csQCA** y **fsQCA**. Además, permite aplicar la mayor parte de los parámetros de ajuste (por ejemplo, en el manejo de los residuales lógicos) y realizar *análisis de necesidad y suficiencia* tanto de las *condiciones individuales* como de las *configuraciones* **(Schneider y Wagemann, 2012, p. 282)**. A nivel gráfico, permite producir XY Plots, histogramas y barcharts. Igualmente, el programa facilita proceder con el análisis de

> *"subconjuntos/superconjuntos, calibrar condiciones difusas, realizar expectativas direccionales para obtener la solución intermedia, así como conseguir las tres soluciones finales con indicación de consistencia y cobertura de los modelos, las configuraciones e incluso de los casos"* **(Schneider y Wagemann, 2012, p. 282).**

Después de una larga temporada sin actualizaciones, muy recientemente, ha aparecido la **fsQCA** versión **3.0** de este software, que por primera vez puede ejecutarse en MacOS además de en Windows.

Otro software utilizado para realizar análisis de **QCA** es **TOSMANA (Cronqvist, 2016)**, que tradicionalmente había permitido ejecutar análisis **csQCA** y **mvQCA**, pero no **fsQCA**. De nuevo, aunque durante mucho tiempo el software no fue actualizado, recientemente ha aparecido la versión **1.52**, que ahora proporciona apoyo preliminar para realizar análisis difuso (*fuzzy*). Como ponen de manifiesto **Schneider y Wagemann (2012, p. 283)**, se trata de un software particularmente útil para obtener la visualización de los datos a través de diagramas de Venn y producir cálculos Booleanos. En realidad, tal como afirman **Rihoux et al., (2009, p. 171)**,

fsQCA y **TOSMANA** ofrecen herramientas complementarias en entornos gráficos sencillos para el usuario. **TOSMANA** incluye también una utilidad llamada *threshold setter* que permite visualizar los casos para realizar *dicotomizaciones* y *multicotomizaciones* de forma más sencilla y efectiva.

Además, en el proceso de *minimización* permite obtener un listado de los *supuestos simplificadores* usados para la obtención de la solución parsimoniosa **(Schneider y Wagemann, 2012, p. 283)**. Como principales desventajas de **TOSMANA**, **Schneider y Wagemann (2012, pp. 283-284)** recordaban que tradicionalmente no ha venido permitiendo aplicar los parámetros de ajuste, como indicadores de *consistencia y cobertura,* ni tampoco permite realizar expectativas direccionales, cerrando así la posibilidad de obtener *soluciones intermedias*.

STATA y **R** permiten también la aplicación de análisis **QCA**. **Thiem y Dusa (2013)** desarrollaron una versión de **QCA** para **R** que supuso un importante avance en términos de software para *análisis configuracional*. Posteriormente, ambos han desarrollado sus propios paquetes software, como Thiem el **QCApro**, mientras que Dusa, ha desarrollado **QCAGUI**, que permite operar en **R** con un entorno gráfico de usuario. Por su parte, **Medzihorski et al., (2020)** han desarrollado el paquete **SetMethods**. Estos son solo algunos de los ejemplos de herramientas **QCA** para **R** y **STATA**, cada una de las cuales presenta elementos comunes, pero también distintivos en su aproximación al *análisis configuracional*. Dada esta enorme y enriquecedora diversidad de software, la tabla solo indica para estos casos las características más frecuentes o más importantes de estos paquetes. Como puede verse, los entornos **STATA** y **R**, y en particular el segundo, proporcionan hoy en día un elenco de herramientas enorme y en permanente actualización para llevar a cabo análisis configuracional, que da cuenta del estado de efervescencia a nivel metodológico de las técnicas **QCA.**

Cabe mencionar, el gran acervo de software ubicado en **COMPASSS** en el enlace: http://legacy.compasss.org/software.htm



» Software

This page provides an overview of the most common software available for CCMs. All programmes are provided free of charge and their developers have invested much time and effort into developing, programming, maintaining and documenting them. If you make use of software, we would thus ask you to acknowledge this by citing it in your work, also in the interest of scientific transparency and replicability. If you have any questions or comments related to a specific programme, please contact its author(s).

Software	Environment	Methodology
acq	Unix shell	csQCA, fsQCA
FSGoF	DOS	Fit test for fuzzy-set analysis
fs/QCA 2.5	Windows	csQCA, fsQCA
fs/QCA 3.0	Windows, macOS	csQCA, fsQCA
fuzzy	Stata	csQCA, fsQCA
Kirq	Windows, macOS, Linux	csQCA, fsQCA
MDSO/MSDO	Web app	MDSO/MSDO
EvalC3	Microsoft Excel (Windows)	Configurational Prediction Modeling
Tosmana	Windows (.NET/Mono)	csQCA, mvQCA, fsQCA
QCA Add-in for Excel	MS Excel (Windows, Mac)	csQCA, mvQCA, fsQCA
cna	R	Coincidence analysis
SetMethods	R	Additional functions for QCA
LogicOpt	R	Truth table minimization
NCA	R	Necessary conditions analysis
QCA	R	csQCA, mvQCA, fsQCA
QCA3	R	csQCA, mvQCA, fsQCA
QCAfalsePositive	R	Test for Type I errors
QCApro	R	csQCA, mvQCA, fsQCA
QCAtools	R	Additional functions for QCA

Fuente: Compasss (http://legacy.compasss.org/software.htm).

Software fsQCA 3.0

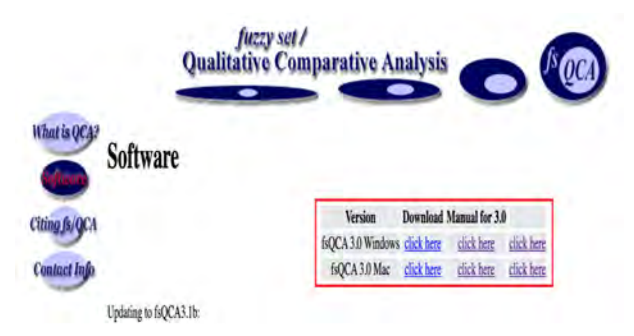
Dadas las facilidades y la popularidad del uso del software cerrado por el Dr. Ragin en su versión **fsQCA 3.0,** a continuación, se presentará desde su localización en el portal, descarga en su laptop y configuración de funciones más importantes, para un mayor entendimiento de su uso.

Configurando el software
Caso 1

Esta sección del capítulo, tiene como objetivo principal, el que conozca cómo configurar el software de análisis cualitativo comparativo de datos difuso o **fsQCA** (*fuzzy-set Qualitative Comparative Analysis*) y que reconozca las principales funciones de las que consta sugiriéndose hacer a través de utilizar el archivo de trabajo: **empresas innovadoras cluster.xls** del caso que se analiza en detalle en el capítulo 4 **(Tabla 4.5)** que se encuentra originalmente en **Excel**.

Para lograrlo, hacer:

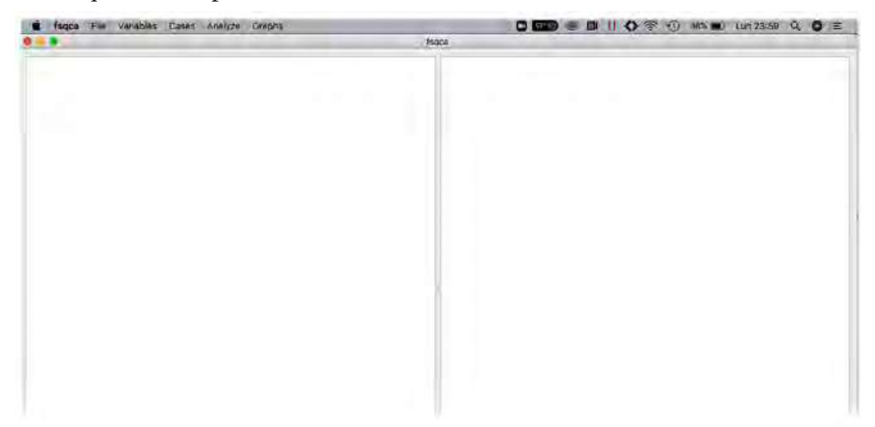
a. Ir a http://www.socsci.uci.edu/~cragin/fsQCA/software.shtml y descargar el software **fsQCA 3.0** de acuerdo con su equipo.



b. Al correr el programa, se deplegará ícono para colocar en su barra de herramientas.

fsqca

c. Al hacer oprimir, observará la siguiente ventana, la cual está dividida en dos partes. La de la izquierda corresponde a la apertura de los archivos y la selección de funciones a aplicarle. La de la derecha, siempre que lo indique, corresponde a la de los resultados.



Abriendo un archivo de datos de varios formatos. Caso 2

Del archivo de trabajo: **empresas innovadoras cluster.csv** caso de la **Tabla 4.5**, que se encuentra en **Excel**, realizar los siguientes procedimientos:

- Del menú, oprima: **File → Open**



- En el cuadro de diálogo, seleccione el archivo que desee abrir y oprima en **Open.**

Los archivos de datos vienen en una amplia variedad de formatos, y el software está diseñado para manejar lo siguiente:

a. Valores separados por comas (*Comma-Separated Values*) (*.csv) o archivo delimitado por comas (*Comma-Delimited File*), hecho en **Excel** u otro software de hoja de cálculo.

b. Separación por espacios (*Space Separated*) (*.txt) o archivo delimitado por espacios, pueden crearse en WORD u otro software, de procesamiento de texto y guardarse solo como texto.

c. Separación por tabulaciones (*Tab Separated*) (*.dat) o archivo delimitado por tabulaciones, pueden ser creados por **SPSS** y otros paquetes de software estadístico.

d. Archivos de datos (*Raw Data Files*) (*.raw) generados por **STATA** (la extensión se puede cambiar a *.csv, si se guarda en el formato delimitado por comas).

e. Los formatos recomendados son * .csv (**Excel**) y * .dat (**SPSS**). Tenga en cuenta, que **fsQCA** realiza los siguientes supuestos sobre la estructura de los archivos de datos * .csv, * .dat y * .txt.:

- Primero, y lo más importante, es que **fsQCA** supone que las celdas en la primera fila de la hoja de cálculo contienen nombres de variables para sus respectivas columnas.
- En segundo lugar, **fsQCA** supone que los datos comienzan en la segunda fila de la hoja de cálculo y que cada caso es una sola fila.
- Finalmente, **fsQCA** supone que cada columna contiene celdas del mismo tipo de datos. Los tipos de datos pueden variar entre columnas, pero deben ser consistentes dentro de las columnas. Recuerde utilizar nombres de variables muy simples, utilizando solo caracteres alfanuméricos sin puntuación o espacios incrustados. Por ejemplo, "JMT1964" está bien, pero "JMT 1964" y "JMT-1964" no lo están.

f. Abrir / Guardar (*Open/Save*), datos creados originalmente en **Excel**: Guarde (*Save*) el archivo de **Excel** en formato * .csv (en formato: valores separados por comas). Asegúrese de que el primera fila de la hoja de cálculo de datos de **Excel**, contenga los nombres de las variables. Abra (*Open*) en **fsQCA.**

g. Abrir / Guardar (*Open/Save*), datos creados originalmente en **SPSS**: Guarde (*Save*) el archivo **SPSS** en formato * .dat (delimitado por tabulaciones) o en formato * .csv (valores separados por comas). **SPSS** le preguntará si desea Escribir nombres de variables en archivo (*Write variable names to file*). No desactive la opción.

h. Abrir / Guardar (*Open/Save*), datos creados originalmente en **STATA**: Guarde (*Save*) el archivo **STATA** en formato * .dta y luego vaya a Archivo, Exportar (*File*, *Export*) y elija el archivo como Datos separados por comas (*Coma-separated data*). En la nueva ventana, inserte el nombre del archivo para Escribir en el archivo (*Write to de file*), luego, para Delimitador (*Delimiter*), elija el formato separado por comas (*Coma-separated format*) y haga clic en Enviar (*Submit*). En algunas versiones de **STATA**, pueden necesitar cambiar el nombre del nuevo archivo * .dta como un archivo * .csv.

i. Abrir / Guardar (*Open/Save*), datos creados originalmente en Word / Bloc de notas (*Notepad*). Ingrese los datos, delimitados por espacios. Asegúrese de que la primera línea, contenga el nombre de las variables, también separados por espacios. Guarde (*Save*) el archivo en formato * .txt (solo texto), TXT (Texto con saltos de línea), TXT (MS-DOS) o TXT (MS-DOS con saltos de línea). Abra (Open) en **fsQCA.**

Opciones para guardar archivos. Caso 3

a. Del menú, oprima: **File->Save as**



b. El archivo de datos modificado se guarda, sobrescribiendo la versión anterior del archivo con el mismo nombre y ubicación, o,

c. Para guardar un nuevo archivo de datos en un diferente formato, desde el menú oprima: **File → Save as**, y:
- El archivo se guardará en formato de valores separados por coma (*Comma-Separated Values*) (*.csv)
- Ingrese (*Enter*) un nombre como archivo de datos nuevo.

Abriendo archivos de otros formatos. Caso 4

Una vez que tenga sus datos en el programa **fsQCA 3.0** y haya completado algunos análisis preliminares, tiene la opción de editar los datos en **fsQCA** (consulte más adelante el apartado: **fsQCA.** Formas de editar) o de hacerlo con la ayuda de otros paquetes de software, con los que pueda estar más familiarizado (por ejemplo, **SPSS** o **Excel**). Del mismo modo, puede mostrar los datos gráficamente con el programa **fsQCA** (consulte el apartado: **fsQCA**. Cómo hacer gráficos) o recurrir a **SPSS**, **STATA** o **Excel** para obtener representaciones gráficas más elaboradas. Si elige **SPSS**, **STATA** o **Excel** para estas operaciones, debe guardar el archivo **fsQCA** y transferirlo al programa que elija.

Excel

Para abrir datos **fsQCA** en **Excel**, guarde la hoja de cálculo de datos fsQCA en valores separados por comas (*Command-Separated Values*) (* .csv). Asegúrese de que el nombre de la variable, en el archivo de datos **fsQCA,** se escriba sin espacios entre ellos (no se permiten espacios incrustados). Así, en **Excel** debe:

- **Oprimir: File → Open…**
- Abrir el archivo **fsQCA** que haya guardado.
- Ahora, puede editar los datos y despegarlos gráficamente en **Excel**.
- Para transferir el archivo **SPSS**, de nuevo a **fsQCA**, consulte: ***fsQCA. Abriendo un archivo de datos de varios formatos***, **punto f.**

SPSS

Para abrir datos **fsQCA** en **SPSS**, guarde la hoja de cálculo de datos **fsQCA** en valores separados por comas (*Comma-Separated Values*) (* .csv) o en un archivo delimitado por comas (*Comma-Delimited File*). Asegúrese de que el nombre de la variable, en el archivo de datos **fsQCA,** se escriba sin espacios entre ellos (no se permiten espacios incrustados). Así, en **SPSS** debe:

- Oprimir: **File → Open → Data…**
- Abrir el archivo **fsQCA** que haya guardado.
- **SPSS** le hará una serie de preguntas acerca del archivo. Cheque cuáles son:

Preguntas que requiere responder en SPSS	Respuestas Sugeridas
Does you text file match a predefined format?	No
How are your variables arranged?	Delimited
Are variable names included at the top of your file?	Yes
Line number that contains variable names	1
What is the decimal symbol?	Period
The first case of data begins with line number?	2
How are your cases represented?	Each line represents a case
How many cases do you want to import?	All of the cases
Which delimiters appear between variables?	Comma
What is the text qualifier?	None
Would you like to save this file format for future use?	Y/N
Would you like to paste the syntax?	No
Then click Finish	

- Ahora, puede editar los datos y mostrarlos gráficamente en **SPSS**.
- Para transferir el archivo **SPSS**, de nuevo a **fsQCA**, consulte: ***fsQCA. Abriendo un archivo de datos de varios formatos, punto g.***

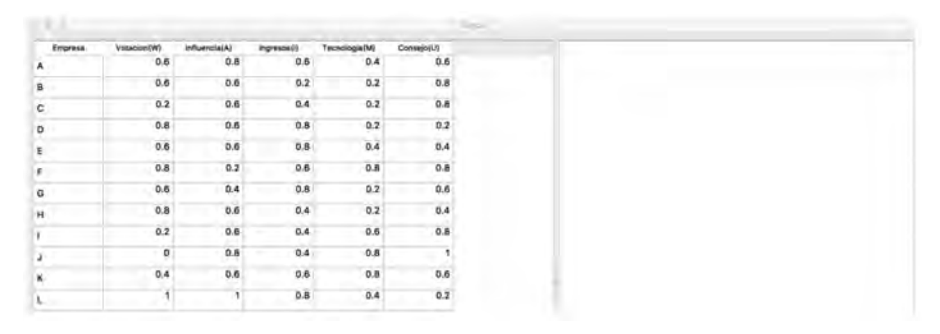
STATA

Para abrir datos **fsQCA** en **STATA**, guarde la hoja de cálculo de datos fsQCA en valores separados por comas (*Command-Separated Values*) (* .csv) o en un archivo delimitado por comas (*Comma-Delimited File*). Asegúrese de que el nombre de la variable, en el archivo de datos **fsQCA,** se escriba sin espacios entre ellos (no se permiten espacios incrustados). Así, en **STATA** debe:

- Oprimir: **File → Import → Text data created by spreadsheet**
- En la nueva ventana, despliegue (*Browse*) para su archivo *.csv y el delimitador (*Delimiter*), elija datos delimitados por coma (*Comma-Delimited Data*).
- Ahora, puede editar y usar los datos en **STATA**.
- Para transferir el archivo **SPSS**, de nuevo a **fsQCA**, consulte: ***fsQCA. Abriendo un archivo de datos de varios formatos, punto h.***

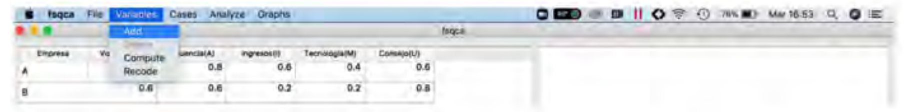
Editando datos de nuevo ingreso.
Caso 5

a. Del menú, oprima: **File → Open → Caso Empresas innovadoras cluster.csv**



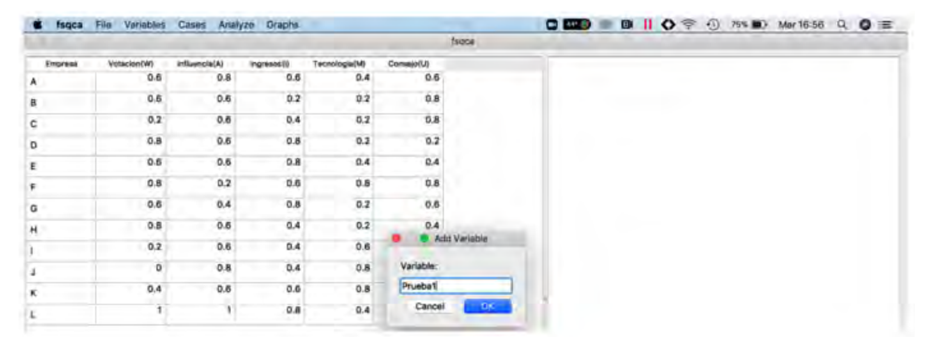
Empresa	Votacion(W)	Influencia(A)	Ingresos(I)	Tecnologia(M)	Consejo(U)
A	0.6	0.8	0.6	0.4	0.6
B	0.6	0.6	0.2	0.2	0.8
C	0.2	0.6	0.4	0.2	0.8
D	0.8	0.6	0.8	0.2	0.2
E	0.6	0.6	0.8	0.4	0.4
F	0.8	0.2	0.6	0.8	0.8
G	0.6	0.4	0.8	0.2	0.6
H	0.8	0.6	0.4	0.2	0.4
I	0.2	0.6	0.4	0.6	0.8
J	0	0.8	0.4	0.8	1
K	0.4	0.6	0.6	0.8	0.6
L	1	1	0.8	0.4	0.2

b. Del menú, oprima: **Variables → Add…**

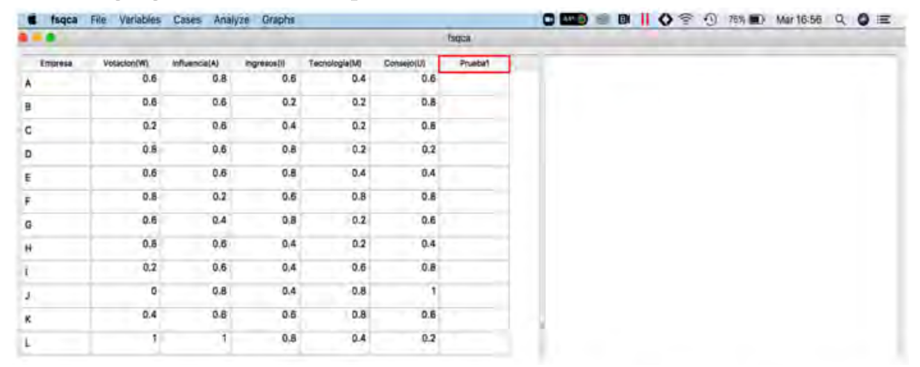


c. La ventana **Add Variable** se abrirá para ingresar el nombre de la variable. Para realizarlo, se debe considerar las siguientes reglas:

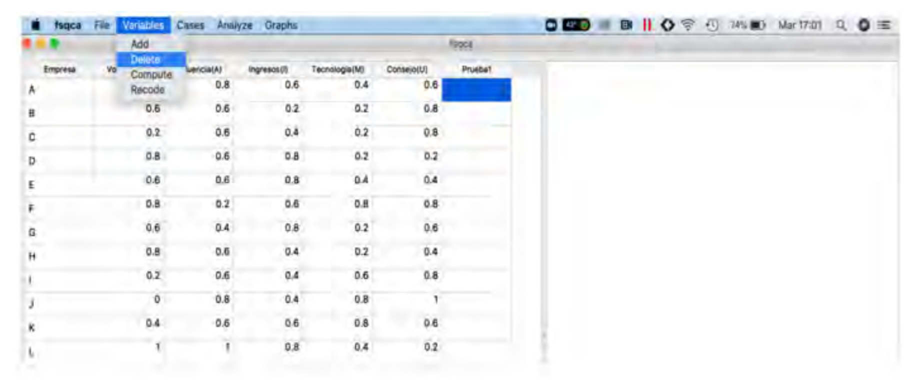
- La longitud del nombre no debe exceder cincuenta caracteres.
- Cada nombre de variable debe ser única; no se permite la duplicidad.
- Los nombres de las variables, no distinguen entre mayúsculas y minúsculas. Los nombres NEWVAR, NewVAR y newvar se consideran idénticos.
- Los nombres de las variables no pueden incluir espacios, guiones o signos de puntuación.
- Solo se pueden usar caracteres alfanuméricos (0-9, a-Z).



- Agregue la variable, oprimiendo el botón **OK.**



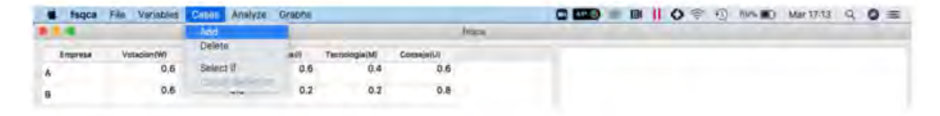
d. También puede borrar variables resaltando la variable y oprimiendo: **Variable → Delete:**



e. Solo basta con confirmar **Yes** para borrar la variable deseada.

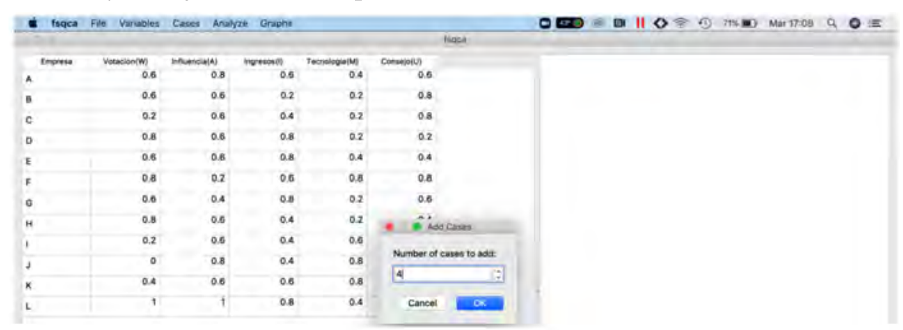


f. Del menú, para agregar casos, simplemente seleccione y oprima: **Cases → Add…**



Observe que, en general, **fsQCA** puede procesar una gran cantidad de casos. Sin embargo, una de sus características principales, es que trata con combinaciones de *condiciones causales*; por lo tanto, agregar más variables influirá en más en tiempo de cómputo que lo que el caso pueda agregar. El número de combinaciones posibles es $\mathbf{2}^{k}$, donde ***k*** es el

número de *condiciones causales*. Como regla general, se recomiendan hasta 10 (es decir, 1024 combinaciones posibles); cuando se trata con más de 10 condiciones, es solo cuestión de la cantidad de tiempo que esté dispuesto a esperar para que el programa haga los análisis. La mayoría de las aplicaciones, usan de tres a ocho *condiciones causales*. Ingrese el número de casos de su conjunto de datos, presione el botón **OK** y la hoja de cálculo aparecerá como:



g. Ingrese los datos. Puede ingresar datos en cualquier orden, ya sea por caso o por variable, para áreas seleccionadas o celdas individuales. La celda activa se resalta con un color más oscuro. Cuando selecciona una celda e ingresa un dato, el valor se muestra en el editor de celdas debajo de la barra de menú. Los valores pueden ser numéricos o de cadena. Los valores de datos no se registran hasta después de presionar **Enter**.

h. Antes de cerrar la hoja de cálculo de datos, guardela (**Save**) como rutina para no perder la información.

Editando datos ya ingresados.
Caso 6

a. Para el caso de ingresar variables que ya existen en la hoja de cálculo de datos, oprima: **Variables → Add.**

b. Ingrese el nombre de la variable y presione el botón **OK.**

c. Para el caso de borrar variables existentes en la hoja de cálculo de datos, resalte la celda en la columna de variable que desee sea borrada. Oprima**:** **Variables → Delete.**

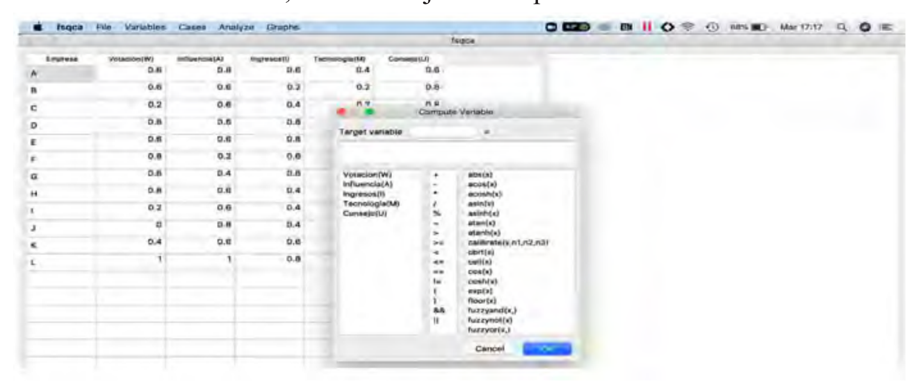
Haciendo cálculos con los datos. Caso 7

Del archivo de trabajo: **empresas innovadoras cluster.csv,** correspondiente al caso de la **Tabla 4.5**, realizar los siguientes procedimientos:

a. Con el fin de hacer cálculos con las nuevas variables a partir de las existentes o expresiones numéricas o lógicas, oprima: **Variables → Compute…**



b. La siguiente ventana se desplegará, con los nombres de las variables en su archivo de datos, listado abajo a la izquierda.



c. Escriba el nombre de una variable objetivo. Puede ser, o una variable existente o una nueva variable, que se agregará al archivo de datos de trabajo. No se recomienda usar una sola letra como variable nombre (por ejemplo, **X**). Esto hará que la función de cálculo se bloquee. Se sugiere seguir las pautas de asignación de nombres de las variables, descritas en la sección: ***fsQCA. Editando datos de nuevo ingreso, en el punto c***.

d. Para construir una expresión, pegue o teclee los componentes dentro del campo Expression (la ventana debajo del nuevo campo variable). Se ofrece el siguiente listado de operadores y funciones que se pueden realizar.

Operadores Aritméticos		
Operador	**Significado**	**Descripción**
+	Suma	El término anterior debe ser agregado al término siguiente. Ambos términos deben ser numéricos
-	Resta	El término siguiente debe ser sustraído del término anterior. Ambos términos deben ser numéricos
*	Multiplicación	Los términos anterior y siguiente están multiplicados. Ambos términos deben ser numéricos
/	División	El término anterior es dividido por el siguiente término. Ambos términos deben ser numéricos y el segundo no debe ser 0
Operadores Relacionales		
Operador	**Significado**	**Descripción**
<	Lógico menor que	Verdadero (= 1) para términos numéricos, si el término anterior es menor que el término siguiente. Es verdadero para los términos de la cadena alfanumérica, si el término anterior aparece antes que el siguiente término en la secuencia de clasificación (en orden alfabético). Este operador normalmente se usa solo en una condición lógica.
>	Lógico mayor que	Verdadero (= 1) para términos numéricos, si el término anterior es mayor que el siguiente. Es verdadero para los términos de la cadena alfanumérica, si el término anterior aparece después que el siguiente en la secuencia de clasificación (en orden alfabético). Este operador normalmente se usa solo en una condición lógica.
<=	Lógico menor o igual que	Verdadero (= 1) para términos numéricos, si el término anterior es menor o igual que el siguiente. Es verdadero para los términos de la cadena alfanumérica, si el término anterior aparece antes que el siguiente en la secuencia de clasificación (en orden alfabético), o si los dos son iguales. Este operador normalmente se usa solo en una condición lógica.

>=	Lógico mayor o igual que	Verdadero (= 1) para términos numéricos, si el término anterior es mayor o igual que el siguiente. Es verdadero para los términos de la cadena alfanumérica, si el término anterior aparece después que el siguiente en la secuencia de clasificación (en orden alfabético), o si los dos son iguales. Este operador normalmente se usa solo en una condición lógica.
=	Lógico igual que	Verdadero (= 1) para términos que son exactamente iguales. Si los términos de la cadena alfanumérica, son de longitud desigual, el término más corto se rellena a la derecha con espacios antes de la comparación. Este operador normalmente se usa, solo en una condición lógica.
<>	Lógico no igual que o desigual que	Verdadero (= 1) para términos que no son exactamente iguales. Si los términos de la cadena alfanumérica, son de longitud desigual, el término más corto se rellena a la derecha con espacios antes de la comparación. Este operador normalmente se usa solo en una condición lógica.
&&	Lógico Y (AND)	Verdadero (= 1) si el término anterior y el siguiente son lógicamente verdaderos. Los términos pueden ser lógicos o numéricos; los términos numéricos mayores que 0, se tratan como verdaderos. Este operador normalmente se usa solo en una condición lógica.
II	Lógico O (OR)	Es verdadero ya sea si el término anterior o el siguiente, son lógicamente verdaderos. Los términos pueden ser lógicos o numéricos; los términos numéricos mayores que 0 se tratan como verdaderos. Este operador normalmente se usa solo en una condición lógica y funciona pegando el símbolo dentro del campo de Expression.
~	Lógico No o negación que	Es verdadero si el siguiente término es falso. 1 - (término numérico). Este operador normalmente se usa solo en una condición lógica.

Funciones Aritméticas		
Operador	**Significado**	**Descripción**
abs (x)	Absoluto	Entrega el valor absoluto de x, el cual debe ser numérico.
acos (x)	Arco Coseno	Entrega el arco coseno (función inversa del coseno) en radianes, el cual debe ser un valor numérico entre 0 y 1.
asin (x)	Arco Seno	Entrega el arco coseno (función inversa del seno) en radianes, el cual debe ser un valor numérico entre 0 y 1.
atan (x)	Arco Tangente	Entrega el arco coseno (función inversa de la tangente) en radianes, el cual debe ser un valor numérico.
ceil (x)	Redondeo hacia arriba	Entrega el valor entero que resulta del redondeo hacia abajo de **x**, la cual debe ser numérica. Por ejemplo, ceil (1.5) = 2.
calibrate	Calibrado	Transforma una escala deintervalo o de razón (*ratio*) dentro de un conjunto difuso (*fuzzy-sets*).
cos (x)	Coseno	Entrega el coseno en radianes, el cual debe ser un valor numérico.
cosh (x)	Coseno Hiperbólico	Entrega el coseno hiperbólico $[(e^{x} + e^{-x}) / 2]$ en radianes y debe ser un valor numérico; **x** no puede exceder el valor de 230.
exp (x)	Exponencial e	Devuelve **e** elevado a la potencia **x**, donde **e** es la base de los logaritmos naturales y **x** es numérico. Valores más grandes de **x** > **230,** producen resultados que exceden la capacidad de la máquina.
floor (x)	Redondeo hacia abajo	Entrega el valor entero que resulta del redondeo hacia abajo de **x**, la cual debe ser numérica. Por ejemplo, ceil (1.5) = 1.
fmod (x,y)	Residuo	Entrega el residuo cuando **x** se divide por el módulo (**y**). Ambos argumentos, deben ser numéricos y el módulo (**y**) no debe ser 0.
fuzzyand (x,…,)	Mínimo difuso	Entrega el mínimo de dos o más conjuntos difusos (*fuzzy-sets*), por ejemplo, fuzzyand (1.0,0.1)= 0.1

fuzzyor (x,…,)	Máximo difuso	Entrega el máximo de dos o más conjuntos difusos (*fuzzy-sets*), por ejemplo, fuzzyand (1.0,0.1)= 1.0
fuzzynot (x)	Negación difuso	Entrega la negación (1-**x**) de conjuntos difusos (igual al No lógico ~), por ejemplo, fuzzynot (0.8)= 0.2
int (x)	Entero	Entrega la parte entera de parte de **x.** Los números son redondeados hacia abajo al entero más cercano.
log (x)	Logaritmo base e	Entrega el logaritmo de **x** con base **e**, el cual debe ser numérico y más grande de 0
log 10 (x)	Logaritmo base 10	Entrega el logaritmo de **x** con base **10**, el cual debe ser numérico y más grande de 0
pow (x,y)	Potencia	Devuelve el término anterior elevado a la potencia del siguiente término. Si el término anterior es negativo, el siguiente término debe ser un número entero. Este operador puede producir valores demasiado grandes o demasiado pequeños para que la computadora los procese, particularmente si el siguiente término (el exponente) es muy grande o muy pequeño.
round (x)	Redondeo	Entrega el entero que resulta de redondear **x**, el cual, debe ser numérico. Los números que terminan en **0.5**, exactamente se redondean desde 0. Por ejemplo: round (2.5) = 3.0
sen (x)	Seno	Entrega el seno en radianes, el cual debe ser un valor numérico.
senh (x)	Seno Hiperbólico	Entrega el seno hiperbólico $[(e^x - e^x) / 2]$ en radianes y debe ser un valor numérico; **x** no puede exceder el valor de 230
square (x)	Cuadrado	Entrega el cuadrado de **x,** el cual debe ser numérico.
sqrt (x)	Raíz Cuadrada	Entrega la raíz cuadrada de **x,** el cual debe ser numérico y no negativo.
tan (x)	Tangente	Entrega el el valor de la relación seno/coseno que es el de la tangente, en radianes por lo que es numérico.
tanh (x)	Tangente Hiperbólica	Entrega el seno hiperbólico $[(e^x - e^{-x}) / (e^x + e^{-x})]$ en radianes y debe ser un valor numérico.

Otros Operadores		
Operador	**Significado**	**Descripción**
( )	Agrupación	Los operadores y las funciones entre paréntesis, se evalúan antes que los operadores y las funciones fuera de los paréntesis.
”	Marca	Se usa para indicar los valores de las variables de cadena alfanumérica.Por ejemplo: Compute if….: Variable == “NA”.
SYSMIS	Subconjunto	Se usa para seleccionar subconjuntos de casos. Por ejemplo: Select if ...: Variable == SYSMISS.
Clear	Borrar	Borra el texto en el campo **Expression.**

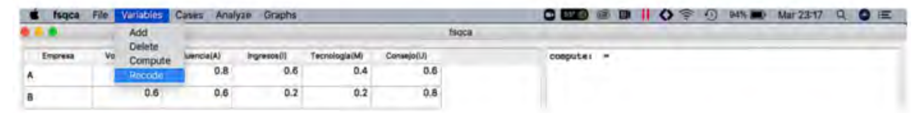
Recodificando dentro de las mismas variables.
Caso 8

Del archivo de trabajo: **empresas innovadoras cluster.csv,** caso de la **Tabla 4.5**, realizar los siguientes procedimientos:

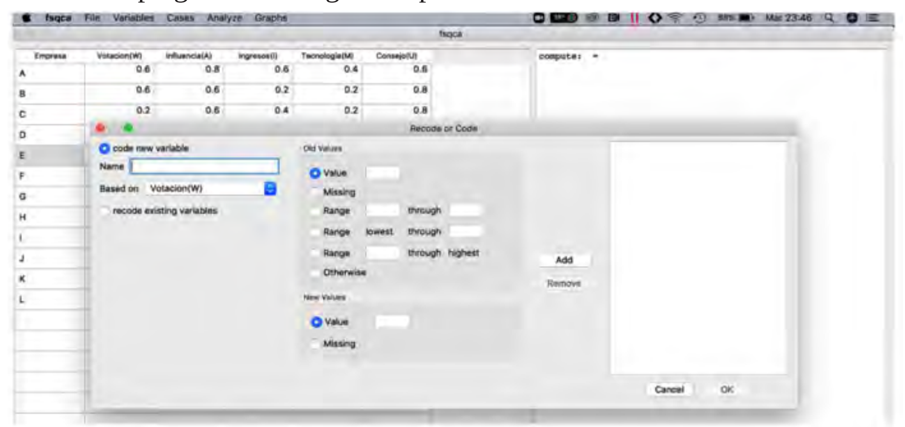
a. Puede modificar los datos recodificándolos. Esto es particularmente útil para contraer o combinar categorías.

b. Puede recodificar los valores dentro de las variables existentes, o puede crear nuevas variables, basadas en los valores registrados de las variables existentes. Recodificar dentro de las mismas variables, reasigna los valores de las variables existentes o contrae los rangos de los valores existentes, en nuevos valores. Puede recodificar variables numéricas y de cadena alfanumérica; puede recodificar variables simples o múltiples; no tienen que ser del mismo tipo. Puede recodificar variables numéricas y de cadena alfanumérica juntas. Así:

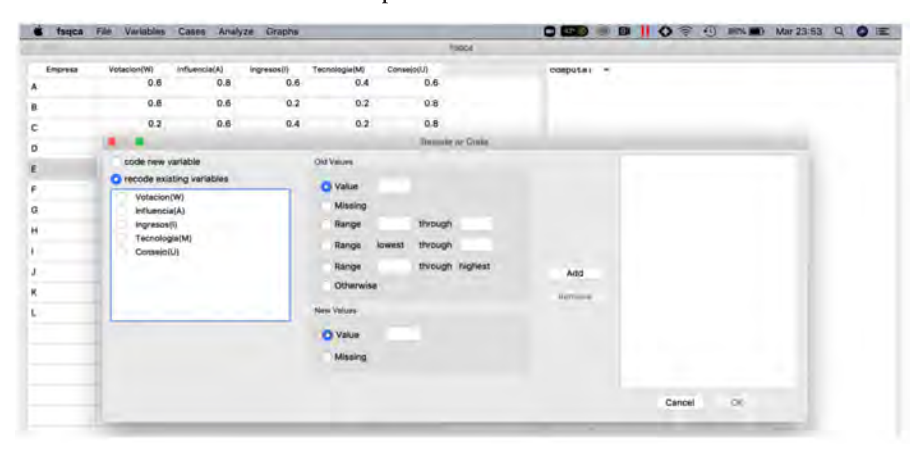
I. Para realizar la recodificación de los valores de una variable, oprima: **Variables → Recode**



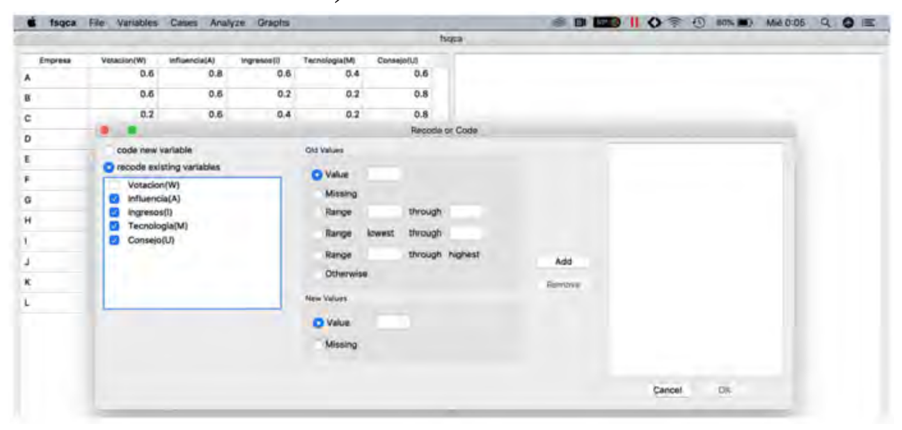
II. Desplegándose la siguiente pantalla



III. Seleccione la opción **recode existing variables**, una ventana con las variables existentes aparecerá



IV. Seleccione las variables que requiera recodificar (numéricas o de cadena alfanuméricas)



V. Puede definir los valores a recodificar, usando las ventanas **Old Values** y **New Value**

Ventana	Descripción
Old Value(s)	Puede recodificar valores individuales, rangos de valores y valores faltantes. Los rangos no pueden seleccionarse para variables de cadena alfanumérica. Los rangos incluyen, sus puntos finales y cualquier valor perdido por el usuario. que se encuentre dentro del rango.
New Value	El único valor dentro del cual, cada viejo valor o rango de valores, es recodificado. Puede ingresar un valor o asignar el valor faltante.

VI. Ingrese sus especificaciones a la lista de la derecha.

Recodificando dentro de diferentes variables. Caso 9

Recodificar dentro de diferentes variables, reasigna los valores de variables existentes o contrae rangos de valores existentes, dentro de nuevos valores para una nueva variable. Se pueden:

- Recodificar variables numéricas y de cadena alfanumérica.
- Recodificar variables numéricas dentro de variables de cadenas alfanuméricas y viceversa.

A fin de recodificar los valores de **Old Variable dentro de New Variable,** se oprime: **Variables → Recode…**

- Seleccione **code new variable** así como las variables existentes que requiera recodificar del menú.
- Ingrese un nombre de variable de salida (nuevo). Especifique como recodificar los valores.

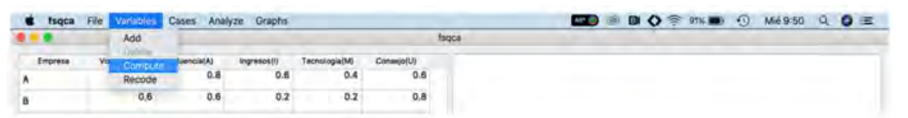
Calibrando el conjunto de datos difusos (*fuzzy-sets*) Caso 10

Del archivo de trabajo: **empresas innovadoras cluster.csv,** caso de la **Tabla 4.5**, realizar los siguientes procedimientos:

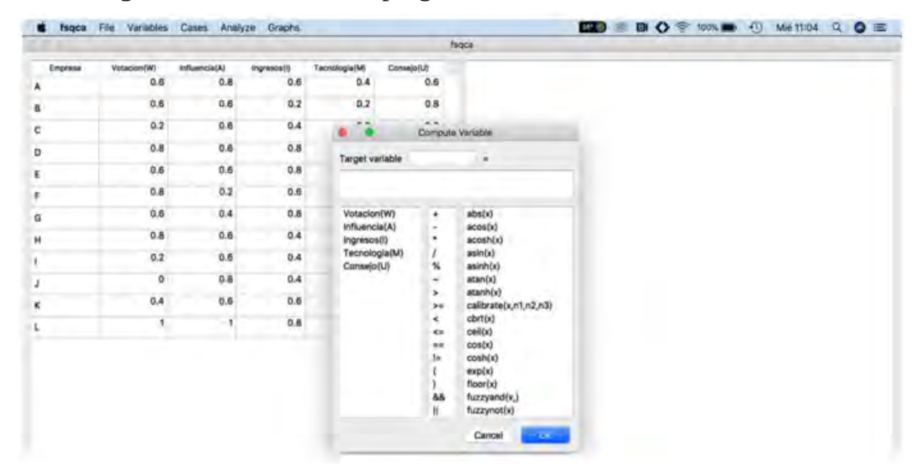
Para transformar las convencionales variables de escala de intervalo y razón en conjuntos de *datos difusos*, es necesario calibrarlas, de modo que las variables coincidan o se ajusten a los estándares externos. La mayoría de los científicos sociales no lo toma en cuenta y utiliza mediciones no calibradas, es decir, muestran de manera simple, las posiciones de los casos entre sí. Sin embargo, las medidas no calibradas son claramente inferiores a las medidas calibradas. Por ejemplo, con una medida de innovación sin calibrar es posible saber que una organización o país hace más actividades de innovación que otro o más innovación que el promedio, pero aún no se sabe si sigue una tendencia de surgimiento de innovaciones cerrada o si su grado de novedad es más radical que incremental.

Los conjuntos de *datos difusos*, se calibran utilizando criterios teóricos así como empíricos sustantivos externos a los datos, y tienen en cuenta la conceptualización, definición y etiquetado del conjunto en cuestión, por parte del investigador. El producto final, es la calibración detallada del grado de pertenencia de los casos en conjuntos difusos, con puntajes que varían de **0.0** a **1.0.** El investigador, debe especificar los valores de una variable de escala de intervalo o de razón que corresponden a tres puntos de referencia cualitativos que estructuran un conjunto difuso: el umbral para la pertenencia completa (puntaje difuso = **0.95**), el umbral para la no pertenencia completa (puntaje difuso = **0.05**) y punto de cruce (puntaje difuso = **0.5**). Estos tres puntos de referencia, se utilizan para transformar la proporción original o los valores de escala de intervalo en puntajes de pertenencia difusos, utilizando transformaciones basadas en las probabilidades de registro de la pertenencia completa. Así:

a. Del menú, oprima: **Variables → Compute**

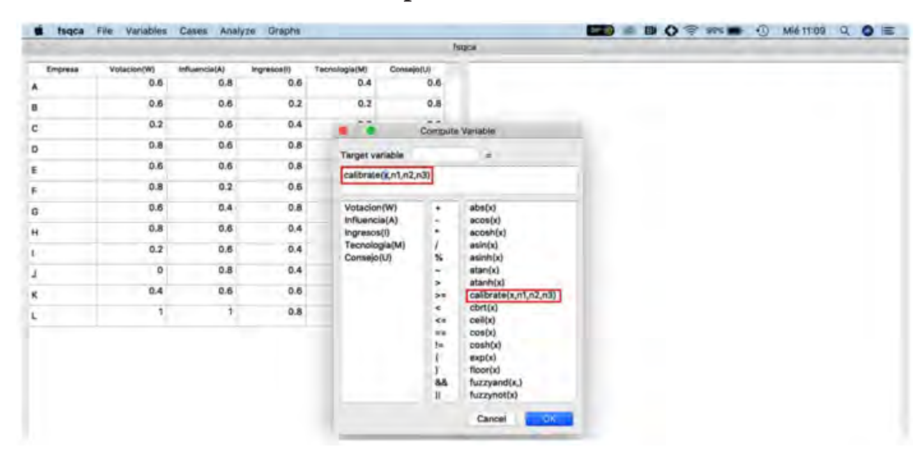


b. La siguiente ventana se desplegará:



c. Nombre la nueva variable (usando de **2-8** caracteres estándar sin espacios, guiones o puntuación) para el conjunto de datos difuso.

d. Oprima el botón **calibrate (x, n1, n2, n3)** en el menú de funciones, que lo transferirá a la ventana **Expressions**



e. Edite la expresión **calibrate (,,,),** por ejemplo, **calibrate (tecnologia, 25,10,2)**. Aquí, **tecnologia** es el nombre de la variable de escala de intervalo o razón existente del archivo, que puede transferir desde el menú **Variables** a la izquierda. El primer número es el valor de **oldvar** que corresponde al umbral de pertenencia completa en el conjunto de objetivos (**0.95**), el segundo número es el valor de **oldvar** que corresponde al punto de cruce (**0.5**) en el conjunto de objetivos y el tercero número es el valor de **oldvar** que corresponde al umbral para la no pertenencia completa en el conjunto objetivo (**0.05**). Oprima **OK.**

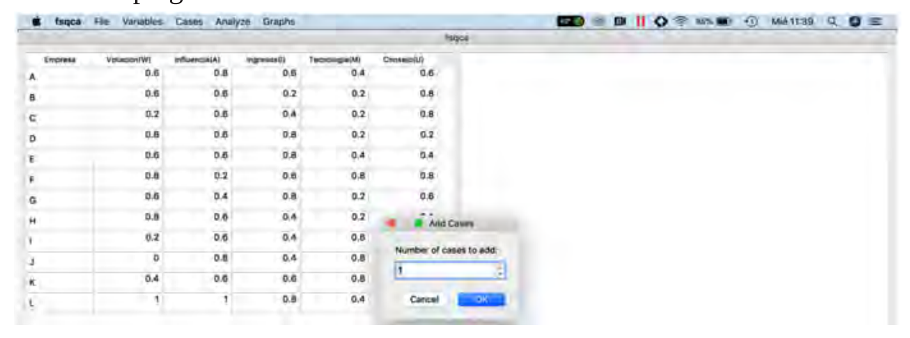
f. Verifique la hoja de cálculo de datos para asegurarse de que los puntajes difusos, corresponden a los valores originales de la manera prevista. Puede ser útil ordenar la variable en orden descendente o ascendente, haciendo clic en el nombre de la variable en el encabezado de la columna. El resultado es una calibración detallada del grado de pertenencia de los casos en conjuntos, con puntajes que van de **(0) a (1)**.

Editando casos de inserción.
Caso 11

Del archivo de trabajo: **empresas innovadoras cluster.csv,** caso de la **Tabla 4.5**, realizar los siguientes procedimientos:

a. Para agregar casos, a una hoja de datos ya existente, oprima: **Cases → Add…**

b. Se desplegará la ventana:



c. Ingrese el número de casos que desea agregar. Los casos adicionales aparecerán al final (abajo) de la hoja de cálculo de datos.

Editando casos a borrar.
Caso 12

a. Para borrar casos individuales de una hoja de datos ya existente, oprima: **Cases → Delete,** la siguiente ventana aparecerá:



b. Con esta función solo puede eliminar un caso a la vez.

c. **fsQCA**, le preguntará si desea eliminar el caso en el que ha resaltado una celda en la hoja de datos.

Editando casos de selección condicional Si (*If*). Caso 13

Los casos de selección **If** (si) proporciona varios métodos para seleccionar un subgrupo de casos, basados en criterios que incluyen variables y expresiones complejas, como:

- Valores de variables y rangos
- Expresiones aritméticas
- Expresiones lógicas
- Funciones

Los casos no seleccionados, permanecen en el archivo de datos pero se excluyen del análisis. Así:

a. A fin de seleccionar un subconjunto de casos para análisis, oprima: **Cases → Select**

b. La siguiente ventana se desplegará:



c. Especifique los criterios para seleccionar casos.

d. Si el resultado de una expresión condicional es verdadera, se selecciona el caso. Si el resultado de una expresión condicional, es falso, no se selecciona el caso.

e. La mayoría de las expresiones condicionales usan una o más de los seis operadores relacionales **(<, >, <=, > =, ==, !, =).**

f. Las expresiones condicionales, pueden incluir nombres de variables, constantes, operadores aritméticos, funciones numéricas y de otro tipo, variables lógicas y operadores relacionales. Seleccionar **If** (Si) funciona mejor cuando es univariante. Por ejemplo, si desea usar la función Seleccionar **If** (Si) que combina dos enunciados lógicos, por ejemplo, un **AND** lógico y un **NOT** lógico, intente crear una nueva variable (con cálculo **Compute** o recodificación **Recode**), que refleje sus criterios de selección y luego use la nueva variable con Seleccionar **If** (Si).

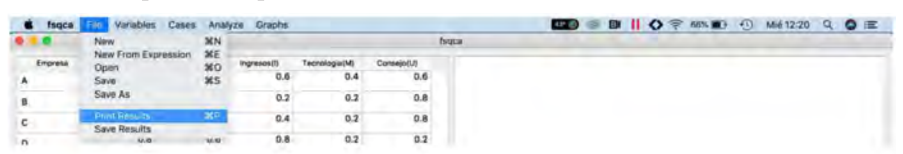
g. Si desea revertir su selección, oprima: **Cases → Cancel Selection...**

Trabajando con salidas de respuesta. Caso 14

Del archivo de trabajo: **empresas innovadoras cluster.csv,** caso de la **Tabla 4.5**, realizar los siguientes procedimientos:

Cuando ejecuta un procedimiento, los resultados se muestran en la ventana de **fsQCA**. Puede usar el desplazamiento hacia arriba y hacia abajo de la ventana para examinar los resultados. Así:

a. Para impresión, oprima: **File→Print Results**



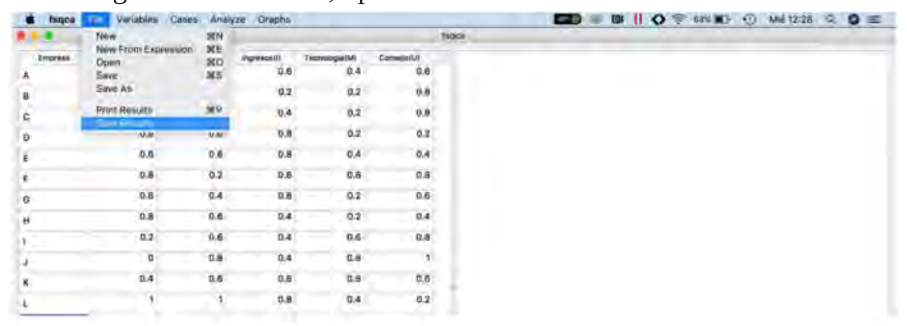
b. Aparecerá la ventana de opciones de la impresora específica de su computadora, en la que puede especificar sus opciones de impresión.



c. La salida se escribe en monoespacio New Courier (10) para permitir el intercambio simple entre programas. Por lo tanto, si abre el archivo * .out en SPSS o en algún otro programa, los números en las tablas se desplazarán ligeramente, a menos que especifique la fuente adecuada.

d. La salida también se puede copiar y pegar en Word, Wordpad, Texto u otro tipo de archivo.

e. Para guardar resultados, oprima: **File → Save Results**



f. **fsQCA** guardará los resultados en formato * .txt (texto sin formato).



Análisis de Condiciones Necesarias (*Necessary Conditions*). Caso 15

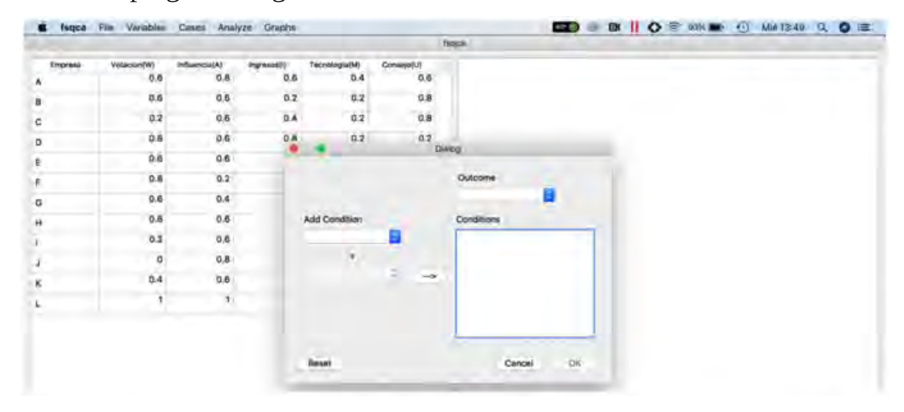
Del archivo de trabajo: **empresas innovadoras cluster.csv,** caso de la **Tabla 4.5**, realizar los siguientes procedimientos:

El análisis de condiciones necesarias (**Necessary Conditions**), produce puntuaciones de *consistencia* y *cobertura*, para condiciones individuales y / o condiciones sustituibles especificadas.

a. Para analizar las condiciones necesarias (**Necessary Conditions**), oprima: **Analyze → Necessary Conditions…**



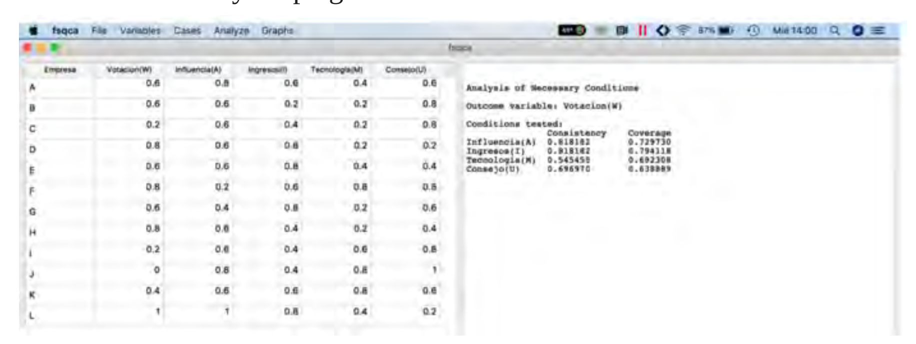
b. Se desplegará la siguiente ventana:



c. Seleccione el resultado a partir del menú en su parte inferior. Posteriormente, seleccione una condición del menú desplegable **Add Conditions** y después transfiérala al cuadro **Conditions** en el lado derecho de la ventana de diálogo. Puede especificar, condiciones necesarias sustituibles utilizando el lógico o, (**OR**, +):



d. Una vez que ha ingresado las especificaciones, presione **OK** y el análisis será realizado y desplegado:



En este contexto, la *consistencia* indica el grado en que la *condición causal*, es un superconjunto del resultado; la *cobertura* indica la relevancia empírica de un superconjunto consistente. **QCA** se basa en análisis de *necesidad y de suficiencia*. Una *condición explicativa* es necesaria si está presente siempre que se produce el *resultado de interés* (aunque en algunos casos, en que la condición está presente el resultado de interés no se produzca), y es *suficiente* si el resultado de interés se produce siempre que la condición está presente (aunque pueda producirse también en ausencia de dicha condición) **(Rihoux y Ragin, 2009, p. xix)**. En este apartado, haremos un paréntesis en nuestra explicación para mencionar las herramientas que **QCA** nos ofrece para analizar relaciones de necesidad. En

particular, el programa **fsQCA 3.0** ofrece una herramienta de análisis de *condiciones necesarias* que muestra medidas de *consistencia* y *cobertura* de la necesidad. Siguiendo con el ejemplo de la base de datos anterior, al realizar dicho análisis observamos lo que muestra la imagen, describiendo:

a. La *consistencia* (*consistency*) en el análisis de *condiciones necesidad* (*analysis of necessary conditions*) indica el *grado de necesidad de una condición*, o dicho de otra forma, *indica la proporción de casos que tienen tanto una condición como el resultado de interés de entre el total de casos que muestran dicho resultado de interés* **(Bol y Luppi, 2013, p. 206)**. En nuestro caso, por ejemplo, la consistencia de la necesidad de la condición Influencia **A** es **0.818182**, lo que indica que **A,** está presente en casi el **82%** de los casos en que el *resultado* (detallado en el *outcome variable*) está presente *votación* **W.** No es el caso, pero por otro lado, si apareciera en los resultados adicionalmente la ausencia de **A** y/o **I** y/o **M** y/o **U** se diría que son necesarias en la totalidad (y/o parcialidad) de los casos, en los que ocurriera.

b. En el análisis de *condiciones necesidad* (*analysis of necessary conditions*) aparece también el concepto de *cobertura* (*coverage*), que indica la *proporción de casos en los que aparecen tanto la condición como el resultado de interés de entre los casos que muestran dicha condición* **(Bol y Luppi, 2013, p. 206)**. Por ejemplo, en el análisis de la imagen anterior, la condición *influencia* **A** tendría una *cobertura* de **0.729730**, lo que indica que en el **73%** de los casos en los que dicha condición está presente, también lo está el resultado de interés *votación,* **W.** Lo más alto registrado como caso, es el de *ingresos* **I,** con cobertura de **0.794118** que significa que el **79%** de los casos en los que dicha condición está presente, también lo está el resultado de interés *votación,* **W.**

Aunque el análisis de *suficiencia* sea una pieza central de los estudios que utilizan **QCA**, lo cierto es que se ha reivindicado el potencial analítico del análisis de necesidad y se han sugerido nuevas estrategias para explorar su potencial **(Bol y Luppi, 2013)**. Una de las razones que explican la rele-

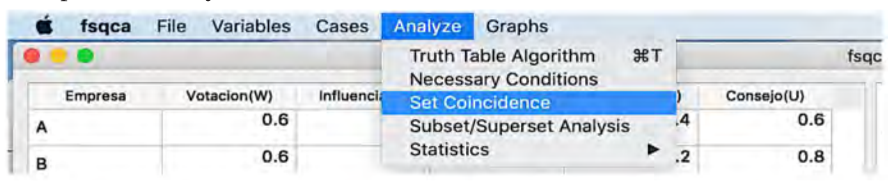
vancia de este análisis es la capacidad de *matizar algunas interpretaciones precipitadas* que se pueden derivar del análisis de *condiciones suficientes* como, por ejemplo, la *existencia de falsas condiciones necesarias. Una falsa condición necesaria* es aquella que, basada en la suficiencia de la solución final, es considerada, *precipitadamente y sin acierto*, como *necesaria* por el solo hecho de estar presente tanto en las diversas rutas como en las distintas soluciones, aunque no resulte ser tal al hacer el *análisis de necesidad (analysis of necessary conditions)* **(Schneider y Wagemann, 2012, p. 226)**.

Análisis de Coincidencia de Conjunto (*Set Coincident Analysis*). Caso 16

Del archivo de trabajo: **empresas innovadoras cluster.csv,** caso de la **Tabla 4.5**, realizar los siguientes procedimientos:

El análisis de coincidencia de conjuntos (**Set Coincidence**) evalúa el grado de superposición de dos o más conjuntos. Así:

a. Para realizar el análisis de coincidencia de conjuntos (**Set Coincidence**), oprima**: Analyze → Set Coincidence**



b. Se desplegará la siguiente ventana



c. Seleccione las condiciones que desea evaluar. Por ejemplo, puede seleccionar todas las condiciones sin resultado, para evaluar el grado de superposición de todas las combinaciones posibles.



d. Una vez que haya ingresado las especificaciones, presione **OK** y el análisis será realizado y desplegado

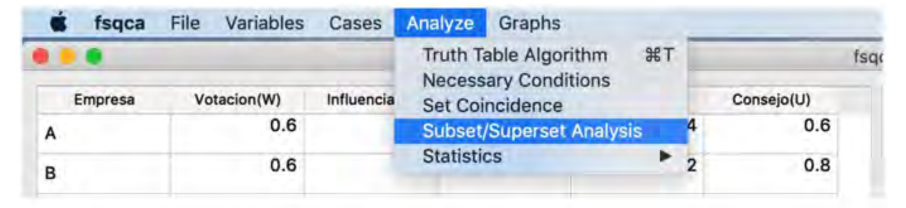
```
File: /Users/DCA/Desktop/Caso W.csv
coincidence(Influencia(A),Ingresos(I),Tecnologia(M),Consejo(U)) = 0.367347
coincidence(Influencia(A),Ingresos(I),Tecnologia(M)) = 0.422222
coincidence(Influencia(A),Tecnologia(M),Consejo(U)) = 0.456522
coincidence(Ingresos(I),Tecnologia(M),Consejo(U)) = 0.434783
coincidence(Influencia(A),Ingresos(I),Consejo(U)) = 0.458333
coincidence(Influencia(A),Tecnologia(M)) = 0.536585
coincidence(Influencia(A),Ingresos(I)) = 0.651163
coincidence(Ingresos(I),Tecnologia(M)) = 0.538462
coincidence(Tecnologia(M),Consejo(U)) = 0.631579
coincidence(Influencia(A),Consejo(U)) = 0.622222
coincidence(Ingresos(I),Consejo(U)) = 0.555556
```

Análisis de subconjunto / superconjunto (*Subset / Superset Analysis*). Caso 17

Del archivo de trabajo: **empresas innovadoras cluster.csv,** caso de la **Tabla 4.5**, realizar los siguientes procedimientos:

El análisis de subconjunto / superconjunto (**Subset/Superset Analysis**), proporciona puntajes de *consistencia* y *cobertura*, para *condiciones* y *configuraciones de condiciones*, así como un puntaje combinado (que es experimental). Proporciona una forma de examinar la el *ser suficiente* de una trayectoria causal (solución) hipotética, así como todos los subconjuntos de condiciones de dicha trayectoria causal. Así:

a. Para realizar el análisis de subconjunto / superconjunto (**Subset/ Superset Analysis**), oprima: **Analyze → Subset/Superset Analysis**



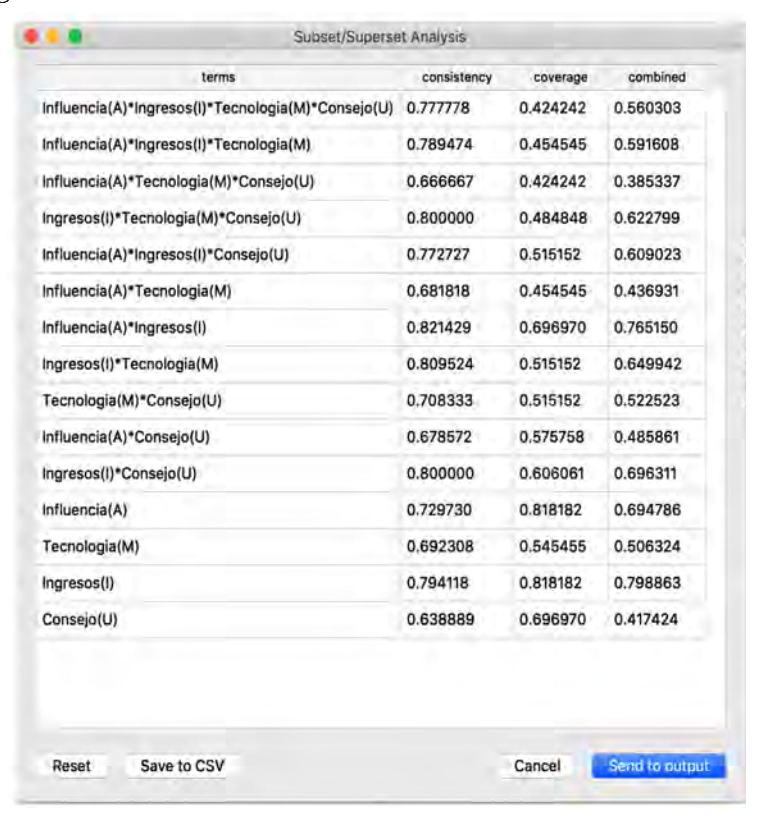
b. Se desplegará la siguiente ventana:



c. Seleccione la variable de resultado y oprima **Set**. Posteriormente, elija las *condiciones causales* y haga oprima **Add** o **Add Negated**, según sus expectativas.



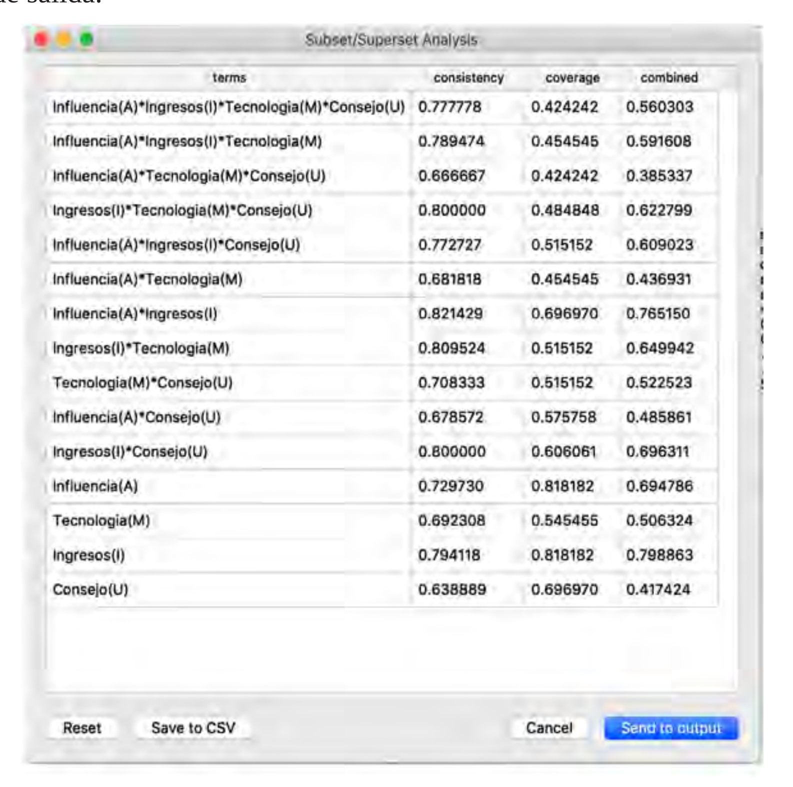
d. Una vez ingresadas las especificaciones, oprima **OK** y le aparecerá la siguiente ventana:


Subset/Superset Analysis

terms	consistency	coverage	combined
Influencia(A)*Ingresos(I)*Tecnologia(M)*Consejo(U)	0.777778	0.424242	0.560303
Influencia(A)*Ingresos(I)*Tecnologia(M)	0.789474	0.454545	0.591608
Influencia(A)*Tecnologia(M)*Consejo(U)	0.666667	0.424242	0.385337
Ingresos(I)*Tecnologia(M)*Consejo(U)	0.800000	0.484848	0.622799
Influencia(A)*Ingresos(I)*Consejo(U)	0.772727	0.515152	0.609023
Influencia(A)*Tecnologia(M)	0.681818	0.454545	0.436931
Influencia(A)*Ingresos(I)	0.821429	0.696970	0.765150
Ingresos(I)*Tecnologia(M)	0.809524	0.515152	0.649942
Tecnologia(M)*Consejo(U)	0.708333	0.515152	0.522523
Influencia(A)*Consejo(U)	0.678572	0.575758	0.485861
Ingresos(I)*Consejo(U)	0.800000	0.606061	0.696311
Influencia(A)	0.729730	0.818182	0.694786
Tecnologia(M)	0.692308	0.545455	0.506324
Ingresos(I)	0.794118	0.818182	0.798863
Consejo(U)	0.638889	0.696970	0.417424

Reset | Save to CSV | Cancel | Send to output

e. Una vez que haya ejecutado el análisis, puede elegir guardar los resultados en un archivo en formato * .csv o enviar el resultado a la ventana de salida.


Subset/Superset Analysis

terms	consistency	coverage	combined
Influencia(A)*Ingresos(I)*Tecnologia(M)*Consejo(U)	0.777778	0.424242	0.560303
Influencia(A)*Ingresos(I)*Tecnologia(M)	0.789474	0.454545	0.591608
Influencia(A)*Tecnologia(M)*Consejo(U)	0.666667	0.424242	0.385337
Ingresos(I)*Tecnologia(M)*Consejo(U)	0.800000	0.484848	0.622799
Influencia(A)*Ingresos(I)*Consejo(U)	0.772727	0.515152	0.609023
Influencia(A)*Tecnologia(M)	0.681818	0.454545	0.436931
Influencia(A)*Ingresos(I)	0.821429	0.696970	0.765150
Ingresos(I)*Tecnologia(M)	0.809524	0.515152	0.649942
Tecnologia(M)*Consejo(U)	0.708333	0.515152	0.522523
Influencia(A)*Consejo(U)	0.678572	0.575758	0.485861
Ingresos(I)*Consejo(U)	0.800000	0.606061	0.696311
Influencia(A)	0.729730	0.818182	0.694786
Tecnologia(M)	0.692308	0.545455	0.506324
Ingresos(I)	0.794118	0.818182	0.798863
Consejo(U)	0.638889	0.696970	0.417424

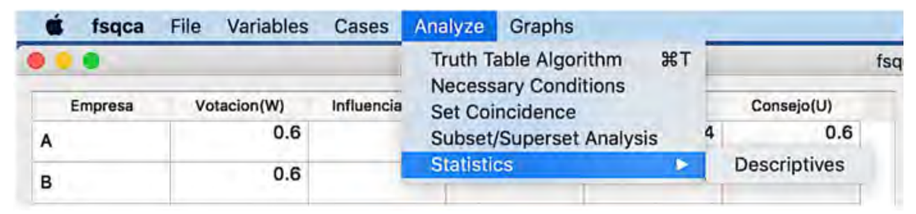
Reset Save to CSV Cancel Send to output

Análisis estadístico descriptivo (*Descriptives*). Caso 18

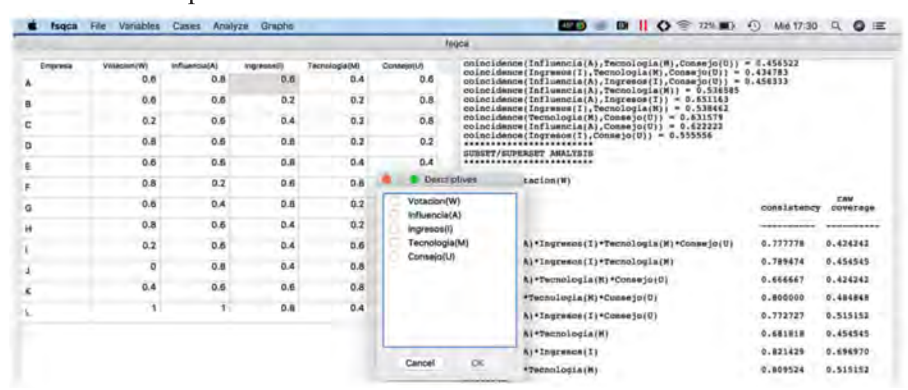
Del archivo de trabajo: **empresas innovadoras cluster.csv,** caso de la **Tabla 4.5**, realizar los siguientes procedimientos:

El análisis estadístico descriptivo (**Descriptives**) muestra un resumen de las técnicas estadísticas univariadas, para condiciones específicas en una sola tabla. Así:

a. Para realizar el análisis estadístico descriptivo (**Descriptives**), oprima: **Analyze → Statistics…**



b. Seleccione una o más condiciones desde la columna **Variables** y transfiéralos . Oprima **OK**



c. La ventana de salida o resultados, mostrará los estadísticos descriptivos correspondientes

Variable	Mean	Std. Dev.	Minimum	Maximum	N Cases	Missing
Votacion(W)	0.55	0.284312	0	1	12	0
Influencia(A)	0.6166667	0.1907587	0.2	1	12	0
Ingresos(I)	0.5666667	0.1972027	0.2	0.8	12	0
Tecnologia(M)	0.4333333	0.2426703	0.2	0.8	12	0
Consejo(U)	0.6	0.244949	0.2	1	12	0

La primera línea, indicará el procedimiento que ha elegido de estadística descriptiva). Las columnas en la tabla descriptiva indican lo siguiente:

Concepto	Abreviatura
La variable seleccionada	Variable
El valor promedio	Mean
La desviación estándar	Std. Dev.
El valor más bajo de la variable	Minimum
El valor más alto de la variable	Maximum
El número de casos	N cases
El número de casos perdidos	Missing

Gráficas.
Caso 19

Del archivo de trabajo: **empresas innovadoras cluster.csv,** caso de la **Tabla 4.5**, realizar los siguientes procedimientos:

Los diferentes programas de **QCA** permiten la realización de una serie de gráficos que en ocasiones pueden servir para ilustrar de forma intuitiva los análisis. Uno de los más sencillo es la graficación de los datos en **fsQCA**, el cual se realiza:

a. Oprimir: **Graphs → XY Plot**



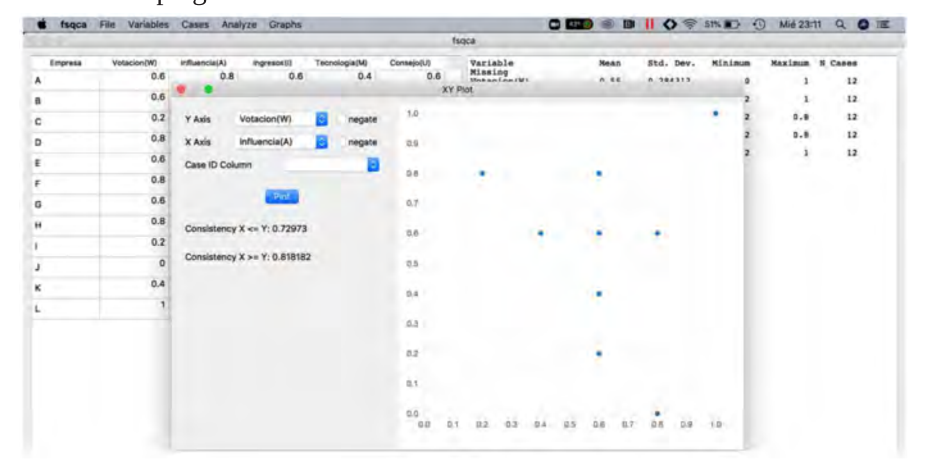
b. Se desplegará, la ventana:



c. Seleccione una variable a definir los valores del eje **X** así como de **Y**, que se muestran en el gráfico.

d. También, puede agregar más información eligiendo una variable de ID de caso (**Case ID Variable**). Esta variable no se representará en el gráfico, pero puede determinar su valor moviendo el cursor a un punto particular en el gráfico, después de haberlo trazado. Por ejemplo, la variable de ID de caso (**Case ID**) podría ser una variable de cadena alfanumérica con los nombres de las organizaciones, empresas, países, etcétera en el conjunto de datos. Una vez graficado, es posible mover el cursor a cualquier punto del mismo, y aparecerá una ventana con el nombre del caso y los valores **x** e **y** del punto.

e. Una vez ingresado las especificaciones, oprima el botón **Plot** y la gráfica será desplegada.



f. Los números debajo del botón **Plot**, muestran la puntuación de la *consistencia* del marco teórica de la investigación. La línea superior, muestra el grado en que los datos trazados son consistentes con **(X) <= (Y)** (**X** es un subconjunto de **Y**). La línea inferior. muestra el grado en que los datos trazados son consistentes con **(X) >= (Y)** (**Y** es un subconjunto de **X**). Si uno de estos dos números indica alta *consistencia*, el otro puede interpretarse como un puntaje de *cobertura*. Por ejemplo, si el número en la línea superior es **0.91** y el número en la línea inferior es

0.63, estos cálculos indican que los datos son en gran medida consistentes con el argumento de que **X** es un subconjunto de **Y** y su *cobertura* de **Y** es **63 %** Es decir, **X** representa el **63%** de la suma de las perte-nencias en **Y**.

g. Puede negar variables en el gráfico, haciendo clic en la opción negar al lado del nombre de la variable. Esta característica, restará el valor del conjunto de datos difuso de esta variable de **(1)**. Ejemplo: Desigualdad = **0.4**; negación de la desigualdad = **0.6**. (Igual que **'~'** y '**fuzzynot (X)**'). Puede copiar el gráfico como una imagen y pegarlo en Word, Texto u otros archivos.

CAPÍTULO 3.

Análisis Cualitativo Comparativo de Datos Nítidos (csQCA)

Esta parte, se refiere al análisis de datos sociales orientados a la innovación, de forma dicotómica, que reflejan la pertenencia de casos en conjuntos convencionales y nítidos. Se pueden encontrar discusiones en profundidad de este método en **Ragin (1987, p. 2000)**. La estrategia de datos analítica, utilizada aquí, se conoce como análisis cualitativo comparativo (**QCA.** *Qualitative Comparative Analysis*). **QCA** se basa en el álgebra Booleana, donde un caso está dentro o fuera de un conjunto, y **QCA** usa datos codificados en binario, con **(1)** indicando pertenencia y **(0)** indi-cando no pertenencia. El **QCA** que utiliza conjuntos convencionales y nítidos, también se conoce como **csQCA**. En este capítulo se ofrecen las nociones básicas de la lógica con la que opera **QCA** en su variante *crisp-set* **QCA** (**csQCA**, de aquí en adelante), mediante la introducción de dos secciones: por un lado, sus fundamentos y aplicaciones y, por otro lado, una guía práctica de cómo usar el software **fsQCA 3.0** con datos hipotéticos, basados en investigacipnes previas.

Conceptos básicos

La base conceptual subyacente de **csQCA** es el álgebra Booleana, la cual fue desarrollada por el matemático George Boole durante el siglo XIX **(Rihoux y De Meur, 2009)**. En **csQCA** se hace una distinción binaria de la realidad donde los casos pueden ser miembros o no de un conjunto. Así, el valor **(1)** indica *pertenencia* en un conjunto y el valor **(0)** indica *no pertenencia*, o *presencia y ausencia*, respectivamente. Se trata, de la primera técnica desarrollada por el Dr. Charles Ragin en 1980 y por el programador Kriss Drass **(Rihoux y De Meur, 2009)** y, hasta la fecha, esta variante sigue siendo la

técnica más utilizada por los investigadores, sin demerito del desarrollo e innovación de las otras variantes.

El proceso de análisis de **csQCA** incluye diversas etapas que han sido desarrolladas desde la publicación de la obra ya mencionada de Charles Ragin en 1987. Así, asumiendo que hay *diversidad* tanto en los *casos* como en las *condiciones*, y que un criterio sistemático de selección ha sido aplicado, el procedimiento consiste en la *dicotomización* de las condiciones, la creación de la *tabla dicotómica*, el *análisis de necesidad*, la construcción de la *tabla de verdad*, el *análisis de suficiencia*, el proceso de *minimización*, así como la *evaluación de los resultados* **(Rihoux y Ragin, 2009; Wagemann y Schneider, 2007, p. 2012)**. A continuación se presenta cada una de las etapas del análisis de **csQCA** con ejemplos ilustrativos para su mayor comprensión.

Uso de datos binarios

Hay dos condiciones o estados en el álgebra booleana: verdadero (*o presente*) y falso (*o ausente*). Estos dos estados están representados en la base 2: **(1)** indica presencia; **(0)** indica ausencia. El análisis comparativo típico basado en Boole aborda la presencia / ausencia de condiciones bajo las cuales se obtiene cierto resultado. Por lo tanto, en un análisis booleano de datos sociales, todas las variables, *condiciones causales* y resultados deben ser medidas de escala nominal, preferiblemente binarias. Las medidas de escala de intervalo, se transforman en medidas de escala nominal de múltiples categorías. Las medidas de escala nominal, con más de dos categorías se representan con varias variables binarias.

Negación Booleana

En la lógica booleana, la negación cambia los puntajes de pertenecia de **(1 a 0)** y de **(0 a 1)**. Por ejemplo, la negación del conjunto de datos nítidos de empresas innovadoras, es el conjunto de datos nítido de empresas no empresas innovadoras. Si un caso, tiene una puntuación booleana de **(1)** en el conjunto de empresas innovadoras, entonces, tiene una puntuación booleana de **(0)** en el conjunto de empresas no innovadoras.

Atribución de valores y calibración

La atribución de valores es siempre una operación delicada en ciencias sociales, en especial en lo que corresponde a la administración de la innovación. Mucho más en **QCA**, pues al operar con una *lógica determinista* y habitualmente un *número mediano de casos*, las técnicas *configuracionales* son muy sensibles a mínimas variaciones en los casos o en los valores asignados a ellos de la matriz de datos **(De Meur et al., 2009, p. 155)**. En principio, existen fenómenos sociales para los que la atribución de valores es *a priori* sencilla. Por ejemplo, categorías *dicotómicas* (dentro-fuera) se prestan a una construcción poco controvertida de condiciones de naturaleza nítida. En otros casos, sin embargo, debemos elegir puntos de corte en fenómenos que tienen una naturaleza más *continua* (o al menos ordinal), al construir *condiciones dicotómicas o multicotómicas*, por ejemplo, si se requiere *dicotomizar* una condición referida a la edad de un CEO. En estos casos, los investigadores deben ser transparentes al justificar explícitamente, preferiblemente con base en argumentos teóricos o empíricos, los puntos de *dicotomización* (o multicotomización) elegidos **(Rihoux y De Meur·, 2009, p. 42)**.

A diferencia de las condiciones nítidas o multicotómicas, la atribución de valores para condiciones difusas, se basa en la idea de calibración. *La calibración* implica la identificación de puntos concretos dentro de los valores continuos (u ordinales o de intervalos) de una variable en los que se producen cambios cualitativos para un fenómeno **(Ragin, 2009, p. 90)**. La identificación de estos puntos de referencia, conocidos como *anclajes o referencias*, permite *transformar una variable no calibrada en una condición difusa* que asigne grados de pertenencia a un conjunto. Por ejemplo, reflexionemos ahora con más detenimiento en la edad de un CEO y su importancia en la innovación en el ámbito de las ciencias sociales. Un abordaje puramente cuantitativo a la edad, atribuiría para este factor valores entre **0 y 120** (años), aproximadamente, esperando que las variaciones en la edad produzcan diferentes comportamientos y preferencias. Sin embargo, sabemos que no todos los valores cuantitativos de la edad tienen el mismo significado cualitativo; por ejemplo, sabemos que a los **18** años, en México,

se establece la mayoría de edad y, con ello, el ejercicio del derecho al voto. De igual forma, se puede establecer que a los **65** años de edad, se accede a la jubilación laboral. Por lo tanto, se puede recurrir a *criterios cualitativos* para identificar que los **18** años es la edad de referencia para indicar que una persona es *mayor de edad* y que los **65** años marcan la edad de *jubilación* . Planteando una condición que considere la pertenencia de diversos individuos al conjunto *personas ancianas/mayores*, se podría fijar una *anclaje o referencia* a los **67** años, que indicará total pertenencia y otro a los **18** años que fije la *total exclusión* a dicho conjunto. Se debería encontrar una edad intermedia entre estas dos edades para fijar el punto de *máxima indefinición, que no siempre tiene que ser la media o la mediana estadística* que, pese a ser un recurso fácil, puede carecer de sentido teórico, cosa fundamental para las investigaciones **QCA**.

Por lo tanto, las condiciones difusas nos permiten tener en cuenta estos *umbrales cualitativos* a través de la *calibración*. *La calibración*, según **Ragin (2007, p. 10)**, *la calibración* permite unir lo mejor de la investigación cuantitativa y cualitativa. Idealmente, la selección de puntos de *calibración* debería también estar basados en el conocimiento empírico sustantivo del investigador y justificarse explícitamente **(Ragin, 2007, p. 11)**. Las investigaciones **QCA** deben dedicar al menos unas frases, normalmente en secciones metodológicas, a la discusión y justificación transparente de estas *decisiones de calibración* **(Schneider y Wagemann, 2010, p. 7). Ragin (2007, p. 12)** propone dos métodos para calibrar condiciones difusas:

1. *El método directo de calibración y el método indirecto de calibración*. Estos dos métodos son aceptados por igual en investigaciones **QCA**, pese a que el primero es más habitual. El *método directo de* calibración, cuya aplicación práctica es más orientada **fsQCA**, consiste en la selección de tres umbrales cualitativos, *anclajes o referencias teóricas*, para calibrar un fenómeno: *el punto de total inclusión en el conjunto, el punto de máxima indefinición de un conjunto a otro, y el punto de total exclusión del conjunto*. Con el *método directo*, el investigador decide qué

puntos dentro de una serie de valores numéricos marcan cambios cualitativos, indicando los puntos en los que los casos estarían *totalmente incluidos* en un determinado fenómeno estudiado (como, por ejemplo, ser un *CEO maduro o mayor*), *más incluidos que excluidos de dicho fenómeno, o totalmente excluidos del mismo*.

2. *El método indirecto* estima grados de pertenencia a partir de agrupaciones semiautomáticas de los casos.
3. Debe tenerse en cuenta, por último, que **Thiem y Dusa (2013, p. 54)** proponen una tercera manera de calibración: el *método de calibración cualitativa* también llamado *método directo de atribución*, en el que el investigador atribuye valores basándose solo en su conocimiento teórico o empírico sustantivo de los casos. Este método puede ser útil cuando no existe información previa que facilite la calibración, siendo el investigador quien decide qué valor se le asigna a cada caso.

Uso de tablas de verdad para representar los datos

Aunque las matrices de datos son un primer paso esencial para iniciar un análisis **QCA**, con esta técnica, se llevan a cabo las *minimizaciones* a través de una *tabla de verdad* que a su vez permiten llevar a cabo una evaluación y tratamiento de posibles *configuraciones contradictorias*.

Para utilizar el álgebra booleana como técnica de comparación cualitativa, es necesario construir una matriz de datos sin procesar, como una *tabla de verdad*. La idea detrás de una *tabla de verdad* es simple. Una vez que los datos se han recodificado en variables de escala nominal y se han representado en forma binaria (como **1** y **0**), solo es necesario clasificar los datos en sus diferentes combinaciones de valores, en condiciones casuales. Cada combinación lógica de valores, en las *condiciones causales*, se representa como una fila de la tabla de verdad. Una vez que se construye esta parte de la *tabla de verdad*, a cada fila se le asigna un valor de salida (una puntuación de **1** o **0** en el resultado), en función de las puntuaciones de los casos que comparten esa combinación de valores de entrada (esa combinación de puntuaciones en la relación de las *condiciones causales*). Así, tanto

las diferentes combinaciones de valores de entrada (*condiciones causales*), como sus valores de salida asociados (el resultado) se resumen en una *tabla de verdad*.

Construcción de la tabla de verdad

A diferencia de la *tabla de datos dicotómicos*, donde cada fila denota un determinado caso o unidad de observación, las filas en las *tablas de verdad* denotan *configuraciones o combinación de condiciones con el* **AND** *Lógico*. Las *tablas de verdad*, tienen tantas filas como combinaciones lógicamente posibles de valores, en las *condiciones causales*. Si hay tres *condiciones causales* binarias, por ejemplo, la *tabla de verdad* contendrá $\mathbf{2^3 =}$ 8 filas, una para cada combinación lógicamente posible de tres condiciones, de presencia / ausencia. La *tabla de verdad*, para un conjunto de datos de tamaño moderado, con tres condiciones binarias y un resultado binario (con **1= presente** y **0= ausente**), se muestra en la **Tabla 3.1.** Técnicamente, no hay razón para incluir la frecuencia de cada combinación, como parte de la tabla de la verdad. Estos valores se incluyen en los ejemplos, para recordarle al lector que cada fila no es un solo caso, sino un resumen de todos los casos con una cierta combinación de valores de entrada. A este respecto, una fila de una *tabla de verdad*, es como una celda de una clasificación cruzada de varias vías de varias variables categóricas independientes.

Para analizar desde las *perspectivas de caso* y *configuraciones* es la **Tabla 3.1**, que se muestra una *tabla de verdad* hipotética, sobre el modelo de innovación abierta de **Mejía-Trejo (2019d)**.

Tabla 3.1. Que muestra una tabla de verdad hipotética, sobre el modelo de innovación abierta de Mejía-Trejo (2019d)

Items	Columnas					
	Condiciones			Resultado		Casos
	KMG	OBM	IEC	OIN	(~OIN)	
1	1	1	1	0	1	A, J
2	0	0	1	0	1	B, E
3	0	1	1	0	1	C
4	0	1	0	1	0	D
5	1	1	0	1	0	F
6	1	0	1	1	0	G
7	1	0	0	1	0	H
8	0	0	0	1	0	1

Nota: donde el signo (~) denota nula, carencia o bajo nivel del factor correspondiente.

Fuente: Mejía-Trejo (2019d).

De la que:

1. Se observan **2**^{3}, siendo *tres el número de condiciones y ocho el número final de filas*. Dado que más de un caso puede tener la misma configuración de valores para cada condición, cada fila de la *tabla de verdad* puede incluir varios casos. Se ha incluido además de la presencia del resultado, su negación **(~ OIN).**
2. *Desde una perspectiva de casos*, el resultado con valor **(1)** indica que aquellos casos con determinadas concliciones también muestran el resultado de interés. Por ejemplo, en la **fila 1**, se puede observar que los casos **(A, J)** que tienen *una mediana/alta de presencia de prácticas de administración de conocimiento* **KMG**, *adopción de modelos de negocios abiertos* **OBM**, *así como presencia de incorporación a ecosistemas de*

innovación **IEC,** *que presentan nula/baja/ ausencia de efectividad en innovación abierta*, **OIN**.

3. *Desde la perspectiva de las configuraciones* se observa que la conjunción de la **fila 8**, *nula/ escasa/ ausencia/ baja presencia práctica de administración del conocimiento* **KMG**, *nula/ escasa/ ausencia/ baja presencia adopción de modelos de negocios abiertos* **OBM, Y** *nula/ escasa/ ausencia/ baja de incorporación a ecosistemas de innovación* **IEC** es su iciente para la ocurrencia de *mediana/ alta presencia de efectividad de innovación abierta,* **OIN**.

Condiciones nítido (*crisp*), multicotómicas (*multi-value*) y difuso (*fuzzy*)

La presencia o ausencia de condiciones explicativas son conceptos clave en cualquier explicación **QCA**. Presencia o ausencia pueden ser entendidas, en términos de la teoría de conjuntos del marco teórico de la investigación, como *inclusión o exclusión* de un caso en un determinado conjunto **(Ragin, 2009b, p. 88)**. Así, siguiendo el ejemplo de la **Tabla 3.2**

Tabla 3.2. Tabla de verdad hipotética de innovación

Empresa	Columnas				
	Efectividad de la ejecución de la innovación tecnológica	Cambio del director de la ejecución de la innovación tecnológica	Insatisfacción del cliente del producto innovado	Pérdida de posición de mercado	Número de casos (frecuencia)
	E	C	S	M	F
1	0	0	0	0	9
2	1	0	0	1	2
3	0	1	0	1	3
4	0	0	1	1	1
5	1	1	0	1	2
6	1	0	1	1	1
7	0	1	1	1	1
8	1	1	1	1	3

Fuente: Elaboración propia.

En el modelo de *notación clásica* **(Rihoux y De Meur, 2009, p. 34)** la presencia de una condición se expresa usando el nombre de dicha condición en mayúsculas **E,** y la ausencia de una condición se expresa usando el nombre de esa condición en minúscula (**e**). La forma de *notación moderna* es similar, aunque todas las condiciones se escriben en minúscula y para indicar negación de la condición se escribe el signo (~) delante de ella. Si la condición cambio del *director de la ejecución de la innovación tecnológica* **C**, está presente para la Empresa **3**, del caso de la Empresa **3**, puede decirse que pertenece al conjunto de casos en que la condición **C**, está presente. En este ejemplo, la relación entre la condición y el caso es *dicotó-mica* (presencia o ausencia), y las condiciones que operan bajo esta lógica *dicotómica* son denominadas condiciones *nítidas* (*crisp*). Sin embargo, encontramos casos en los que las condiciones pueden ser *multicotómicas* (*multi-value*) o incluso, casos en que la pertenencia puede ser una cuestión de grado o *difusas* (*fuzzy*).

Las condiciones multicotómicas se caracterizan porque a diferencia de las condiciones nítidas, *pueden tomar más de dos valores*. Aunque es teóricamente posible, en estas condiciones multicotómicas no hay por qué presuponer siempre una ordenación o jerarquía entre los valores, que pueden simplemente representar estados o características cualitativamente distintas **(Vink y van Vliet, 2009, p. 269)**. Las multicotomías representan diferentes categorías dentro de una misma condición, que no presencia y ausencia, como cuando se quiere definir la posición de un ingeniero de sistemas dentro de una empresa innovadora (apoyo a la operación interna, apoyo a la operación externa, ventas directas, asistente del CEO) o cuando se hace referencia al estado civil de CEO innovadora (soltero, casado, separado, divorciado, viudo).

Las *condiciones explicativas* que admiten una graduación en la pertenencia del caso en el conjunto se denominan condiciones difusas (*fuzzy*). Estas condiciones indican la pertenencia difusa a un conjunto, o dicho de otra manera, el grado en el que puede decirse de un caso que pertenece a un conjunto **(Ragin 2009, p. 89)**. A diferencia de las condiciones nítidas, en las que los casos solo pueden tener los valores **(0)** o **(1),** en las condiciones

difusas los casos oscilan entre **(0)** y **(1),** siendo **(0.5)** el valor intermedio que indica equidistancia perfecta entre la *pertenencia y la exclusión* en el conjunto. Nótese que los conjuntos difusos *no son variables continuas* al uso. En lugar de ello, *combinan la lógica cuantitativa con el conocimiento cualitativo* de los casos a través del proceso de *calibración*: la especificación del punto concreto a partir del cual puede decirse de un caso que pertenece o no a un determinado conjunto **(Ragin, 2007, 2009, p. 90)**.

Agrupamientos

Así como es posible calcular, el número lógicamente posible de combinaciones ($\mathbf{2}^{k}$), también es posible calcular, el número de agrupaciones lógicamente posibles. La fórmula $\mathbf{3}^{k} \mathbf{- 1}$, donde ***k*,** nuevamente es el número de atributos ($\mathbf{3}^{3} - 1 = 26$). La **Tabla 3.3** muestra las 26 agrupaciones lógicamente posibles de las tres dicotomías presentadas en la **Tabla 3.2.**

Tabla 3.3. Grupos que usan 3 dicotomías

Item	Columnas		
	Configuración inicial (8 combinaciones de 3 factores)	**Agrupamiento que involucran 2 factores**	**Agrupamiento que involucran 1 solo factor**
1	**(E) * (C) * (S)**	**(E) * (C)**	**(E)**
2	**(E) * (C) * (~ S)**	**(E) * (~ C)**	**(~ E)**
3	**(E) * (~C) * (S)**	**(~ E) * (~ C)**	**(C)**
4	**(E) * (~C) * (~ S)**	**(~ E) * (C)**	**(~ C)**
5	**(~ E) * (C) * (S)**	**(E) * (S)**	**(S)**
6	**(~ E) * (~C) * (S)**	**(E) * (~S)**	**(~ S)**
7	**(E) * (C) * (~ S)**	**(~E) * (S)**	—
8	**(~E) * (~C) * (~ S)**	**(~E) * (~S)**	—
9	—	**(C) * (S)**	—
10	—	(C) * (~S)	—
11	—	(~C) * (S)	—
12	—	(~C) * (~S)	—

Fuente: Elaboración propia.

Usando la fórmula que se acaba de describir, las **26** agrupaciones posibles se forman de la siguiente manera: **8** involucran combinaciones de tres factores, **12** involucran combinaciones de dos factores y seis involucran un solo factor.

Adición Boolena

En álgebra Booleana, si (**A + B**) = **Z**, y **A** = 1 y **B** = 1, entonces **Z = 1**. En otras palabras, **(1 + 1) = 1**. La idea básica, en la suma Booleana, es que si alguno de los términos aditivos es satisfecho (presente), entonces el resultado es verdadero (ocurre). La adición, en álgebra Booleana, es equivalente al operador lógico **OR.** (En esta discusión, **OR** mayúscula se usa para indicar **OR** lógico). Por lo tanto, la declaración anterior (**A + B**) = **Z**, se convierte en: si **A** es igual a **(1), OR B** es igual a **(1)**, entonces **Z** es igual a **(1).**

La mejor manera de pensar, en este principio es en términos lógicos, no aritméticos. Considere el colapso de estar fuera de mercado una empresa innovadora; suponga que hay tres condiciones generales que la provocan: efectividad de la ejecución de la innovación tecnológica **E;** cambio del director de la ejecución de la innovación tecnológica **C**; insatisfacción del cliente del producto innovado **S.** La *tabla de verdad*, se encuentra en la **Tabla 3.2** (con **1 = presente y 0 = ausente**). Cada combinación de causas, produce el tener o no una pérdida de *posición de mercado* **M** en diferentes empresas, como se indica en dicha Tabla. (**1= presente; 0= ausente**) Este ejemplo, ilustra sucintamente la naturaleza de la adición Booleana: satis-face cualquiera de las condiciones aditivas y sigue el resultado esperado, además de que cada combinación causal, no produce casos contradictorios. La ecuación Booleana simplificada, es:

$$(M) = (E) + (C) + (S)$$

Y expresa, la relación entre las tres condiciones y la pérdida de posición de mercado de una empresa innovadora, de una manera simple y elegante tanto para instancias negativas como positivas. En pocas palabras: si se

cumple alguna de estas condiciones (o dos o las tres), la empresa perderá posición de mercado.

Multiplicación Boolena

La multiplicación Booleana, difiere sustancialmente de la multiplicación normal. Es relevante porque es la aplicación típica de las ciencias sociales con orientación a la administración de la innovación, del álgebra Booleana refiriéndose al proceso de simplificación de expresiones conocidas como sumas de productos. Un producto, es una combinación particular de *condiciones causales*. Los datos, sobre empresas con pérdida de posicionamiento de mercado, se muestran en la misma **Tabla 3.2** y pueden representarse en forma de sumas de productos (es decir, sin reducción) de la siguiente manera:

(E) * (~C) * (~ S)

+

(~E) * (C) * (~ S)

+

(~ E) * (~C) * (S)

+

(E) * (C) * (~ S)

+

(E) * (~C) * (S)

+

(~ E) * (C) * (S)

+

(E) * (C) * (S)

=

(M)

Cada uno de los siete términos, representa una combinación de *condiciones causales* encontradas en al menos una instancia de la pérdida de posionamiento de mercado de una empresa. Los diferentes términos, son

productos porque representan intersecciones de condiciones (de causas y ausencias de causas). La ecuación, muestra las diferentes combinaciones de *condiciones causales* que están vinculadas a a la pérdida del posicionamiento de mercado.

La multiplicación Booleana, como la suma Booleana, no es aritmética. La expresión **(E) * (~ C) * (~ S)** no significa que el valor de la efectividad de la ejecución de la innovación tecnológica **(E= 1)** se multiplique por el valor del cambio del director de la ejecución de la *innovación tecnológica* **(C= 0)** y por el valor de la insatisfacción del cliente del producto innovado (**S= 0**) para producir un valor resultante de (**M= 0**). Significa simplemente, que una presencia de efectividad *de la ejecución de* la *innovación tecnológica* **E,** se combina con una ausencia cambio del director de la ejecución de la innovación tecnológica **C** y una *ausencia de insatisfacción del cliente del producto innovado,* **S**. La situación total, **(E) * (~ C) * (~ S) = (M)**, ocurre en los datos dos veces. Esta conjunción, carácter de la multiplicación Booleana, da forma a la interpretación de la ecuación de sumas de productos, presentada anteriormente: la *pérdida de posicionamiento de mercado* **M**, ocurre si se obtiene cualquiera de las siete combinaciones de tres causas obtenidas. En álgebra Booleana, la suma indica **OR** lógico y la multiplicación indica **AND** lógico. Las tres *condiciones causales*, se interseccionan (**AND**) de diferentes maneras, para indicar diferentes *con iguraciones* empíricas. Las intersecciones obtenidas, son posteriormente combinadas (**OR**) para conformar ecuaciones de suma productos no reducuidas las cuales describen las diferentes combinaciones de las tres *condiciones causales* vinculadas a la pérdida de *posicionamiento de mercado***, M.**

Lógica combinatoria

El análisis Booleano es combinatorio por diseño. En el análisis de pérdida de *posicionamiento de mercado* **M,** se observa de manera rápida, que tan solo en las primeras cuatro filas de la *tabla de verdad* (**Tabla 3.2**), si alguna de las tres causas está presente, la pérdida de *posicionamiento de mercado*

M, se producirá. Si bien es atractivo tomar este atajo, la ruta tomada por el análisis Booleano es más exigente con los datos. Esto es, porque la ausencia de una causa, tiene el mismo estatus lógico como su presencia en el análisis Booleano. Como se señaló anteriormente, la multiplicación Booleana indica que las condiciones de presencia y ausencia, se combinan, es decir se intersectan.

Considere la segunda fila de la *tabla de verdad* (**Tabla 3.2**), que describe las dos instancias de pérdida de *posicionamiento de mercado* **M,** vinculadas al conflicto de configuración causal **(E) * (~ C) * (~ S) = M**. Una inspección simple, sugiere que en este caso, la pérdida de *posicionamiento de mercado* **M**, resultó de la primera causa: efectividad de la ejecución de la *innovación tecnológica* **E**. Pero tenga en cuenta, que si el investigador tenía información solo en esta fila de la *tabla de verdad*, y no en ninguna de las otras instancias de la pérdida del *posicionamiento de mercado* **M**, se podría concluir que la efectividad de la ejecución de la *innovación tecnológica* **E** causa **M**, solo si el cambio del director de la *ejecución de la innovación tecnológica* **C** y la i*nsatisfacción del cliente del producto innovado* **S,** están ausentes. Esto es lo que indica la combinación de **(E) * (~ C) * (~ S)**. Esta fila por sí sola, no indica si la *efectividad de la ejecución de la innovación tecnológica* **E** causaría la pérdida del *posicionamiento de mercado* **M** en presencia del cambio del director de la *ejecución de la innovación tecnológica* **C**, o la *insatisfacción del cliente del producto innovado* **S** o ambos. Todo lo que el investigador sabe de estos dos casos de **(E) * (~ C) * (~ S)**, es que para que la efectividad de la *ejecución de la innovación tecnológica* **E,** cause la pérdida del *posicionamiento de mercado* **M**, puede ser necesario que las otras condiciones **C** y **S** estén ausentes. Desde una perspectiva Booleana, es completamente factible que, en presencia de una o ambas de otras condiciones, por ejemplo, la configu-ración **(E) * (~C) * (S)**, no se produzca la pérdida del *posicionamiento de mercado* **M**. Para volver a las designaciones originales, puede ser que en presencia de la *insatisfacción del cliente del producto innovado* **S** y la efec-tividad de la *ejecución de la innovación tecnológica* **E** se disipará cuando las dos facciones se unan.

Para impulsar aún más este argumento, suponga que el investigador tenía conocimiento de solo las primeras cuatro filas de la *tabla de verdad*. Los datos deberían respaldar la idea, de que la presencia de cualquiera de las tres condiciones, producen la pérdida del *posicionamiento de mercado* **M,** pero nuevamente, los datos pueden indicar que la efectividad de la *ejecución de la innovación tecnológica* **E** causa la pérdida del *posicionamiento de mercado* **M,** solamente, cuando cambio del director de la ejecución de la *innovación tecnológica* **C** y la *insatisfacción del cliente del producto innovado* **S,** están ausentes **(E * ~ C * ~ S);** cambio del director de la *ejecución de la innovación tecnológica* **C** causa la pérdida del *posicionamiento de mercado* **M,** solamente cuando la efectividad de la *ejecución de la innovación tecnológica* **E** y la *insatisfacción del cliente del producto innovado* **S** están ausentes **(~E * C * ~ S)** y así. Una aplicación estricta de la lógica combinacional, requiere que estas limitaciones sean conclusiones extraídas, de una variedad limitada de casos.

Esta característica de la lógica combinatoria, es consistente con la idea de que los casos, especialmente sus características causalmente relevantes, deban verse de manera integral. El carácter holístico del enfoque Booleano, es consistente con la orientación de los investigadores cualitativos, en las ciencias sociales comparativas que examinan diferentes causas en contexto. Cuando se examina la segunda fila de la *tabla de verdad* (**Tabla 3.2**), no se interpreta como instancias de la pérdida del *posicionamiento de mercado* **M** causadas por la efectividad de la *ejecución de la innovación tecnológica* **E,** sino como instancias de la pérdida del *posicionamiento de mercado* **M** causadas por **(E * ~ C * ~ S).** Por lo tanto, en la comparación basada en el Booleano cualitativo, las causas no se ven de forma aislada sino siempre dentro del contexto de la presencia y ausencia de otras condiciones causalmente relevantes.

Uso de implicaciones principales

Otro concepto Booleano que necesita ser introducido es el concepto de *implicación*. Se dice que una expresión Booleana implica a otra, si la pertenencia del segundo término, es un subconjunto de la pertenencia del primero. Por ejemplo, **A** implica **(A * ~B * ~C)** porque **A** abarca todos los miembros de **(A * ~ B * ~ C)** (es decir, **A** * **~B** * **~C** es un subconjunto de **A**. Este concepto se entiende mejor con un ejemplo. Si **A** indica *empresas económicamente dependientes de las innovaciones disruptivas*, **B** indica la presencia de la *industria electrónica*, y **C** indica *políticas de innovación abierta*, **(A)** abarca a todas las empresas dependientes, mientras que (**A * ~ B * ~C**) abarca todos *empresas económicamente dependientes de las innovaciones disruptivas* **A** que carecen *industria electrónica* **(~B)** y de *políticas de innovación abierta* **(~C)**. Claramente, la pertenencia de **(A * ~B * ~C)** está incluida en la pertenencia de **A.** Por lo tanto, **A** implica **(A * ~B * ~C).**

El concepto de implicación, si bien es obvio, proporciona una herramienta importante para minimizar las expresiones de sumas de productos. Considere la *tabla de verdad* hipotética que se muestra en la **Tabla 3.4**

Tabla 3.4. Tabla de verdad hipotética que muestra tres causas de éxito de un producto innovador (P)

Item	Columna				
	Mercado	Competencia	Fondos	Éxito producto innovador	Frecuencia
	M	C	F	P	
1	1	0	1	1	6
2	0	1	0	1	5
3	1	1	0	1	2
4	1	1	1	1	3
5	1	0	0	0	9
6	0	0	1	0	6
7	0	1	1	0	3
8	0	0	0	0	4

Fuente: Elaboración propia.

Que resume datos sobre tres *condiciones causales* que se cree afectan el éxito de la introducción de un producto innovador con éxito **P,** de una empresa tecnológica en particular de videojuegos:

1. *Un mercado en auge para el producto innovador producido por la empresa,* **M.**
2. La amenaza de la competencia del sector por la simpatía en aliarse C, y
3. *La existencia de fondo de inversión de la competencia* **F,** por contrarrestar el ingreso del producto innovador de la empresa en cuestión.

La ecuación booleana para el éxito del *producto innovador* **P,** que muestra expresiones booleanas no reducidas, es:

(M) * (~C) * (F) + (~M) * (C) * (~F) + (M) * (C) * (~F) + (M) * (C) * (F) = (P)

El primer paso en el análisis Booleano de estos datos, es intentar combinar tantas filas compatibles de la *tabla de verdad*, como sea posible. Tenga en cuenta, que esta parte del proceso de minimización utiliza ilas con un valor de salida de **(1)**, **P** con éxito. Esta primera fase de la minimización de la *tabla de verdad*, produce la siguiente ecuación Booleana parcialmente minimizada, que en efecto convierte una ecuación booleana con cuatro de tres términos variables, en una ecuación con tres términos de dos variables. Ver **Tabla 3.5.**

Tabla 3.5. Ejemplo de minimización.
Caso: dos causas presentes con una causa ausente

Item	Columna		
	Expresión 1	Combina con expresión 2	Para producir
1	M * C * F	M * ~ C * F	M * F
2	M * C * F	M * C * ~ F	M * C
3	M * C* ~ F	~ M * C * ~ F	C * ~ F

Fuente: Elaboración propia.

(M) * (F) + (M) * (C) * + (C) * (~F) = P

Los términos del éxito de la introducción de **P**, son como los de la ecuación anterior y se producen utilizando esta simple regla de minimización: combinar filas que difieren en una sola causa si tienen los mismos valores de salida, los cuales se denominan implicaciones principales. Por lo general, cada implicación principal cubre (es decir, implica) varias expresiones sencillas (filas) de la *tabla de verdad*.

En la ecuación parcialmente minimizada dada anteriormente, por ejemplo, la primera implicación principal: **(M) * (F)** abarca dos expresiones booleanas enumeradas en la *tabla de verdad*: **(M) * (C) * (F) y (M) * (~C) * (F).**

Esta expresión Booleana, parcialmente reducida, ilustra un hallazgo común en el análisis Booleano: a menudo, hay más expresiones reducidas (*implicaciones principales*) que las necesarias para cubrir todas las expresiones sencillas originales. La implicaciones principales de **(M) * (C),** relacionan los términos sencillos **(M) * (C) * (F)** y **(M) * (C) * (~F)**, por ejemplo, estos dos términos sencillos, también están cubiertos por **(M * F)** y **(C * ~ F)**, respectivamente. Por lo tanto, la **(M) * (C)**, puede ser redundante, desde un punto de vista puramente lógico; puede que no sea una implicación principal esencial. Para determinar qué implicaciones principales son lógicamente esenciales, se utiliza un dispositivo de minimización conocido como gráfico de implicaciones principales. La minimización, del gráfico de implicaciones principales, es la segunda fase de la minimización Booleana. Dicho brevemente, el objetivo de esta segunda fase del proceso de minimización, es cubrir tantas expresiones Booleanas sencillas como sea posible, con un número lógicamente mínimo, de implicaciones principales. Este objetivo es producto, de un deseo directo de no generar redundancia. La tabla de implicaciones principales, mapea los enlaces entre las implicaciones principales y las expresiones Booleanas sencillas. Dicha tabla, describe estos enlaces en los datos sobre resultados que se presentan en la **Tabla 3.6.**

Tabla 3.6. Carta de implicaciones principales.
Caso hipotético de causas de éxito de un producto innovador (P)

	Expresiones	**Expresiones Booleanas sencillas**			
		(M) * (C) * (F)	**(M) * (~ C) * (F)**	**(M) * (C) * (~ F)**	**(~M) * (C) * (~ F)**
Implica-ciones Principales	(M) * (F)	X	X	—	—
	(M) * (C)	X	—	X	—
	(C) * (~F)	—	—	X	X

Fuente: Elaboración propia.

Una inspección simple, indica que el número más pequeño de implicaciones principales, necesarios para cubrir todas las expresiones Booleanas sencillas, es dos. Para gráficos de implicaciones principales muy complejos, se necesitan algoritmos informáticos sofisticados. Los implicaciones principales **(M) * (F)** y **(C)** * **(~ F)**, cubren las cuatro expresiones Booleanas sencillas. El análisis de la tabla de implicaciones principales, por lo tanto, conduce a la expresión Booleana reducida final que contiene solo los implicantes primos lógicamente esenciales:

(M) * (F) + (C) * (~ F) = (P)

4. Esta ecuación, establece que la introducción de un *producto innovador con éxito* **P,** ocurre cuando hay un *mercado en auge por el producto innovador* **M**, **AND** un gran *fondo de inversión de la competencia* **(F)** [(**M * F**)] **OR** cuando hay una g*ran amenaza por parte de la simpatía de los competidores de aliarse* (**C**) con un *bajo fondo de inversión de la competencia* (**F**) [**(~ F) (C) * (~ F)**].

Estos simples procedimientos, permiten al investigador estimar una ecuación mínimamente lógica, que describa las diferentes combinaciones de condiciones asociadas con un resultado. La ecuación reducida final muestra las dos combinaciones (lógicamente mínimas) de condiciones que causan **P** exitosos y, por lo tanto, proporciona una declaración explícita de causalidad conjunción múltiple.

Ley de De Morgan

La aplicación de las Ley de De Morgan es sencilla. Considere la solución al caso hipotético de introducción de un *producto innovador con éxito* **P,** presentado anteriormente: **(M) * (F) + (C) * (~ F).** Los elementos que están codificados como presentes en la ecuación reducida (por ejemplo, mercado en el término **(M) * (F)**) se ha recodificado como ausentes, y los elementos que están codificados como ausentes (por ejemplo, **(~ F)** en el término **(C)** * **(~F)**) están recodificados al presente. A continuación, el **AND** lógico se recodifica a **OR** lógico, y el **OR** lógico se recodifica a **AND** lógico. Aplicando estas dos reglas, la expresión:

(M) * (F) + (C) * (~ F) = P

Se convierte, en:

(~M + ~F) * (~C + F) = ~P
(~M * ~C) + (~M * F) + (~C * ~F) = ~P

De acuerdo con esta ecuación, fallará la introducción de un producto innovador con éxito **P**, cuando:

1. **No** (**o es pequeño**) existe un mercado en auge para el producto innovador producido por la empresa **M AND no** existe (**o es pequeño**) amenaza de la competencia del sector por la simpatía en aliarse **C, OR.**
2. **No** (**o es pequeño**) existe un mercado en auge para el *producto innovador producido por la empresa* **M AND** hay la existencia de fondo de inversión de la competencia **F, OR.**
3. **No** (**o es pequeño**) existe amenaza de la competencia del sector *por la simpatía en aliarse* **C AND no (o es pequeña)** hay la existencia de fondo de inversión de la competencia, **F.**

En estas diversas combinaciones, se permiten descubrir relaciones donde es posible prever y explicarse diversos contextos, para la generación de toma de decisiones con soporte.

La Ley De Morgan, produce la negación exacta de una ecuación lógica dada. Si hay combinaciones *residuales* en la *tabla de verdad* y se usan como no importa, entonces los resultados de la aplicación de la Ley De Morgan, arrojarán una declaración lógica que no es lo mismo, que el análisis de la ausencia del resultado. Del mismo modo, si los *residuales* se definen como falsos en el análisis inicial, entonces la aplicación de la Ley de De Morgan a la solución (de los casos positivos), arrojará una declaración lógica que abarque no solo los casos negativos, sino también los *residuales*.

Análisis contrafactual

Los fenómenos sociales en los que la innovación interviene y que ocurren hoy día, se consideran aún como de diversidad limitada, circunstancia que refleja una ventaja, evitar la *complejidad*, ya que de lo contrario, no es posible realizar mediciones. No es fortuito que las categorías de las empresas de innovación, como el prestigio de marca, su posicionamiento de mercado y sus altos ingresos coincidan, tal como no es casualidad, que logren y registren puntuaciones altas en prácticamente todos los indicadores, a nivel nacional y/o mundial tanto de riqueza y bienestar, donde se generalmente se agrupan en la industria de países avanzados.

La diversidad de la innovación y su categorización, tienen múltiples origenes que la acotan no solo por desigualdades de mercado, de regulación, de cultura, sino también de historia. De la historia, es detacable mencionar su reflejo en cómo en los últimos treinta años, los países dependen de una creciente tripolaridad geopolítico-económica, basada en la relación EUA-Rusia-China y que ha provocado fenómenos en los últimos diez años sociales que inciden en la innovación. Particularmente, a pesar de la desaceleración económica de la Unión Europea y los EUA, es notable observar que aún así, siguen siendo polo de atracción de corrientes masivas de migración social perfilándose incluso, el desarrollo de nuevas multietnias en dichas áreas; también, es de considerar el caso reciente de la pandemia de **COVID-19**, que pone en jaque al mundo, en términos de salud y economía y que en menos de cuatro meses del 2020, ha provocado tiempos de crisis

aguda, donde se observa la la introducción de nuevas innovaciones (por ejemplo, en el campo médico: equipo de recolección de enfermos, de pruebas rápidas, de cubrebocas, nuevos protocolos de atención hacia pacientes y sociales nuevas vacunas, fases de control de la misma pandemia, etcétera). Como se aprecia, la historia importa, pero provoca al investigador apreciar una diversidad ilimitada altamente compleja, prefiriendo su contraparte para el estudio de los fenómenos, esto es, la diversidad limitada, la cual se considera que es fundamental para la constitución de los fenómenos de la innovación en el ambiente social, pero que también, complica severamente su análisis.

Uno de los aspectos desafiantes de la *investigación comparativa*, es el hecho de que los investigadores trabajan con muestras de casos **N** relativamente pequeños. Los investigadores a menudo confrontan *más variables que casos*, una situación que es muy complicada, por el hecho de que en la actividad de comparar, generalmente se centran en los aspectos de las combinaciones de los casos que se ajustan *configuracionalmente.*

Por ejemplo, un investigador está interesado en un argumento que especifica una intersección de cuatro *condiciones causales*, del que idealmente, debería considerar las dieciséis combinaciones lógicamente posibles de estas cuatro *condiciones causales*, para proporcionar una evaluación exhaustiva del argumento. Sin embargo, los fenómenos sociales ocurren naturalmente y como se ha establecido, están profundamente limitados en su diversidad. El mundo empírico, casi nunca presenta a los científicos sociales, todas las combinaciones lógicamente posibles de las *condiciones causales relevantes* para respaldar sus argumentos como se muestra en los datos hipotéticos en la **Tabla 3.7**, donde el resultado marcado con **¿?**, el resultado no puede ser determinado.

Tabla 3.7. Tabla de verdad con cuatro *condiciones causales* (A, B, C, D) y una salida (Y)

Item	Columna				
	A	B	C	D	Y
1	No	No	No	No	No
2	No	No	No	Si	¿?
3	No	No	Si	No	¿?
4	No	No	Si	Si	¿?
5	No	Si	No	No	No
6	No	Si	No	Si	No
7	No	Si	Si	No	¿?
8	No	Si	Si	Si	No
9	Si	No	No	No	¿?
10	Si	No	No	Si	¿?
11	Si	No	Si	No	¿?
12	Si	No	Si	Si	¿?
13	Si	Si	No	No	Si
14	Si	Si	No	Si	Si
15	Si	Si	Si	No	¿?
16	Si	Si	Si	Si	¿?

Nota: El resultado Y marcado con ¿?, no puede ser determinado.
Fuente: Elaboración propia.

Si bien la diversidad limitada es fundamental para la constitución de los fenómenos sociales y políticos, también complica gravemente su análisis.

Como un sustituto de las combinaciones ausentes de las *condiciones causales*, los investigadores que realizan la actividad de la comparación, a menudo participan en experimentos de pensamiento, o supuestos **(Weber, 1949)**. Es decir, imaginan casos contrafactuales e hipotetizan sus resultados, utilizando su conocimiento teórico de investigación y sustantivo empírico, para guiar sus evaluaciones. Debido a que **QCA** usa *tablas de verdad* para evaluar patrones de casos cruzados, este proceso de considerar casos contrafactuales (es decir, combinaciones ausentes de *condiciones causales*) es explícito y sistemático. De hecho, esta característica de **QCA** es una de sus fortalezas clave. Sin embargo, la consideración explícita de

casos contrafactuales y la incorporación sistemática de los resultados de tales evaluaciones, en declaraciones sobre patrones de casos cruzados son relativamente nuevos para las ciencias sociales. La especificación de las mejores prácticas con respecto a **QCA** y análisis contrafactual, por lo tanto, es esencial **(Mahoney y Barrenechea, 2017)**.

Otro caso, un investigador postula, basándose en la teoría existente, que las *condiciones causales* **A, B, C y D** están relacionadas de alguna manera con el resultado **Y**. Es decir, es la presencia de estas condiciones, no su ausencia, lo que debería estar relacionado, con la presencia del resultado. La evidencia empírica ,indica que muchos casos de **Y** están asociados con la presencia de *condiciones causales* **A, B y C**, junto con, la ausencia de la condición **D** (es decir, **A·B·C·d = Y**). Sin embargo, el investigador sospecha que todo lo que realmente importa, es tener las tres primeras causas, **A, B y C**. Para que (**A * B * C)** genere **Y**, no es necesario que **D** esté ausente. Sin embargo, no hay casos observados de **A, B y C** combinados con la presencia de **D** (es decir, no hay casos observados de **A * B * C * D**). Por lo tanto, el caso empírico decisivo para determinar si la ausencia de **D** es una parte esencial de la mezcla causal (con **A * B * C**) simplemente no existe.

A través del análisis *contrafactual* (es decir, *un experimento de pensamiento,* o supuesto), el investigador podría declarar, que esta combinación hipotética **(A * B * C * D**) es una instancia probable del resultado **Y.** Es decir, el investigador podría afirmar que **(A * B * C * D),** si existiera, conduciría a **Y.** Este análisis contrafactual, permitiría la siguiente simplificación lógica:

$$\mathbf{(A * B * C * d) + (A * B * C * D) = (Y)}$$

$$\mathbf{(A * B * C) * (d{+}D) = (Y)}$$

$$\mathbf{(A * B * C) = (Y)}$$

¿Cuán factible es esta simplificación? La respuesta a esta pregunta, depende del estado del conocimiento del marco teórico de la investigación y del conocimiento empírico sustantivo relevante que exista sobre la conexión entre **D** e **Y**, en presencia de las otras tres *condiciones causales*

(A * B * C). Si el investigador puede establecer, sobre la base de los conocimientos existentes, que hay muchas razones para esperar que la presencia de **D** contribuya al resultado **Y** en estas condiciones (o, por el contrario, que la ausencia de **D**, no sea un factor contribuyente), entonces el análisis contrafactual presentado es factible. En otras palabras, el conocimiento existente hace que la afirmación de que **(A * B * C * D) = Y** sea un contrafactual fácil, porque implica la adición de una causa redundante **D** a una configuración que se cree que está vinculada al resultado **(A * B * C).**

Una de las fortalezas del **QCA,** es que no solo proporciona herramientas para generar los dos puntos finales del *continuum de complejidad / parsimonia*, sino que también proporciona herramientas para especificar soluciones intermedias. Considere la *tabla de verdad* presentada en la **Tabla 3.7**, que usa **(A, B, C y D)** como *condiciones causales* e **Y** como resultado. Suponga, como antes, que el conocimiento del marco teórico y empírico sustantivo existente, sostiene que es la presencia de estas *condiciones causales*, y no su ausencia, lo que está relacionado con el resultado. Los resultados del análisis, que prohíbe los contrafactuales (es decir, la solución compleja), revela que la combinación (**A * B * c)** explica **Y**. El análisis de esta misma evidencia, permitiendo cualquier contrafactual que produzca un resultado más *parsimonioso* es que **A** por sí solo explica la presencia de **Y**. Concebir estos dos resultados como los dos puntos finales del *continuum de complejidad / parsimonia*, se explica de la siguiente manera:

A * B * c______________________________A

Observe que la solución que privilegia la *complejidad*, **(A * B * c)** es un subconjunto de la solución que privilegia la *parsimonia* **A.** Esto se deduce lógicamente del hecho, de que ambas soluciones deben cubrir las filas de la *tabla de verdad* con **Y** presente; la solución *parsimoniosa* también incorpora algunos de los *residuales*, como casos *contrafactuales* y, por lo tanto, abarca filas adicionales. A lo largo del *continuum de complejidad / parsimonia,* hay otras posibles soluciones a esta misma *tabla de verdad*, por ejemplo, la combinación **(A * B).** Estas soluciones intermedias, se producen

cuando diferentes subconjuntos de los residuos utilizados, para producir la solución *parsimoniosa* se incorporan a los resultados. Estas soluciones intermedias constituyen subconjuntos de la solución más *parsimoniosa* (**A** en este ejemplo) y superconjuntos de la solución que permiten la máxima *complejidad* **(A * B * c).** La relación de subconjunto entre soluciones, se mantiene a lo largo del *continuum de complejidad / parsimonia*. La implicación es que cualquier combinación causal, que use al menos algunas de las *condiciones causales* especificadas en la solución compleja **(A * B * c),** es una solución válida de la *tabla de verdad* siempre que contenga, todas las *condiciones causales* especificadas en la solución *parsimoniosa* **A**. Se deduce entonces, que hay dos soluciones intermedias válidas, para la *tabla de verdad*:

A * B

A * B * c____A * c_________A

Ambas soluciones intermedias **(A * B)** y **(A * c),** son subconjuntos de la solución que privilegia la *parsimonia* y los superconjuntos de la solución, que privilegian la *complejidad*. El primero **(A * B),** permite contrafactuales **(A * B * C * D)** y **(A * B * C * d)** como combinaciones vinculadas al resultado (**Y**). El segundo permite contrafactuales **(A* b * c * D)** y **(A * b * c * d).**

La viabilidad relativa de estas dos soluciones intermedias, depende de la factibilidad de los contrafactuales, que se les han incorporado. Los contrafactuales incorporados, en la primera solución intermedia, son fáciles porque se usan para eliminar (**c**) de la combinación **(A * B * c),** y en este ejemplo, el conocimiento existente respalda la idea de que es la presencia de **C,** no su ausencia, la que está vinculada al resultado **(Y).** Sin embargo, los contrafactuales incorporados en la segunda solución intermedia, son difíciles porque se usan para eliminar **(B)** de **(A * B * c).** Según el conocimiento existente, la presencia de **B** debe estar vinculada a la presencia del resultado **Y.** El principio de que solo se deben incorporar *contrafactuales fáciles*, respalda la selección de **(A * B),** como la *solución intermedia óptima*. Esta solución, es la misma que un investigador convencional orien-

tado a casos, derivaría de esta evidencia, basada en un interés directo en las combinaciones de *condiciones causales* que:

- Comparten los casos positivos (o al menos, un subconjunto de los casos positivos),
- Se cree que están vinculados al resultado, no se muestran, por casos negativos.

Si el mundo empírico, solo presentara a los investigadores casos que solo exhibieran las combinaciones lógicamente posibles de *condiciones causales relevantes*, entonces la investigación social sería mucho más sencilla. Por ejemplo, para casos coincidentes que difieren en una sola *condición causal*, sería posible construir comparaciones estructuradas y enfocadas **(George, 1979)**, lo que a su vez facilitaría enormemente la evaluación de la causalidad.

Es así, que haremos una precisión de los conceptos de los *contrafactuales* y su tratamiento así como de la *diversidad limitada* y los *residuales*. Los *residuales*, se definen en **QCA** como *configuraciones de condiciones* que no tienen casos empíricos reales, es decir, que son *configuraciones que no cuentan con evidencia empírica* **(Ragin y Sonnett, 2004, p. 5)**. Estas configuraciones son formalmente posibles, aunque no siempre teóricamente verosímiles, aunque no hayan ocurrido en la realidad: son por tanto contrafácticos. Estos son el reflejo de la llamada *diversidad limitada* en el método de **QCA** (**Rihoux y Ragin, 2009; Schneider y Wagemann, 2012)**, por lo que no tienen ninguna relación con los llamados valores perdidos o los grados de libertad en los métodos estadísticos.

Suponga otro caso, los datos de la investigación realizada por **Mejía-Trejo (2019c)**, en las que uno de los factores (**DSSMEs**. *Digital Skills in the SMEs*), es desglosado en las variables que están relacionados con las habilidades digitales de las empresas PYME para lograr el modelo de innovación por mercadotecnia digital (**DMIM**. *Digital Marketing Innovation Model*) de la **Tabla 3.8.**

Tabla 3.8. Ejemplo tabla de verdad sobre innovación por mercadotecnia digital DMIMI dependiente del factor DSSME por agrupación de casos

Columna			
DSSMEs			DMIM (Y) Agrupación de casos
STL	DSY	BSY	
1	1	1	A, B, C
1	0	1	D
0	0	0	E, F, G, H
1	0	C(0)	I, J, K, L
1	1	1	M
0	1	0	N
0	1	¿?	Residual
0	0	¿?	Residual

DMIM: Digital Marketing Innovation Model; **DSSMEs:** Digital Skill SMEs; **STL:** Strategic Leadership; **DSY:** Digital Savvy; **BSY:** Business Savvy

Fuente: Mejía-Trejo (2019c) con adaptación propia.

Donde:
DMIM: Modelo de innovación por mercadotecnia digital (*Digital Marketing Innovation Model)*: **Y**
DSSMEs: Habilidades digitales de las Pymes (*Digital Skill SMEs*)
STL: Liderazgo estratégico (*Strategic Leadership*)
DSY: Saber digital (*Digital Savvy*)
BSY: Saber del negocios (*Business Savvy*)

Suponga que estamos interesados en el resultado **DMIM**, los casos que deben ser analizados en el proceso de minimización son **(A, B, C, D y M).** Si se quiere explicar **(~ DMIM),** entonces las configuraciones que interesan son **(E, F, G, H, I, J, K, L y N)** (a pesar de las contradicciones). Hay interrogantes ha resolver:

1. El primer interrogante que surge aquí es: ¿por qué hay *residuales* lógicos en las dos últimas filas? Para comprender la aparición de residuales es necesario explicar ahora con más detalle el concepto de diver-

sidad limitada. **QCA** usa este término para referirse a aquel fenómeno donde configuraciones formalmente posibles no existen en la realidad empírica **(Ragin, 1987, pp. 104-113, 2000; Schneider y Wagemann, 2007a, p. 101),** puesto que no existen casos reales donde dichas *configuraciones* se produjeran. En la *tabla de verdad*, estas *configuraciónes* se representan a través de los *residuales* lógicos **(Schneider y Wagemann, 2012, p. 328)**. Es aquí donde interviene la *diversidad limitada*. Según **Schneider y Wagemann (2012, pp. 154-156)**, el fenómeno de la *diversidad limitada* puede ser explicado por tres razones:

a. *Porque la configuración contradice el sentido común o las leyes de la lógica*, y por lo tanto se crea una paradoja (*residual imposible*), por ejemplo, la presencia de una misma persona en dos lugares a la vez.

b. *Porque aunque la configuración podría existir teóricamente*, *aún no se ha dado empíricamente (agrupación de remanentes),* por ejemplo, un CEO de una empresa latina con liderazgo en fabricación de CPUs.

c. *Porque el número de configuraciones, es mayor que el número de casos (residuales matemáticos),* por ejemplo, si contamos con siete condiciones explicativas para 25 casos. En este caso, se advierten contra el uso de cantidades excesivas de condiciones en los modelos. Usando **csQCA** como referencia para su análisis, los autores afirman que algunos modelos podrían estar *libres de contradicciones,* tan solo por el abuso de condiciones explicativas. Esto es, uno de los criterios de *calidad del análisis configuracional*, consistente en que los modelos deben estar libres de contradicciones, *puede eludirse artificialmente inflando el número de condiciones* introducidas en el modelo. Por ello, el número de condiciones debe, a su juicio, mantenerse en una cifra baja en relación al número de casos. En el artículo citado podrá encontrar el lector unas tablas detalladas con la cantidad de condiciones recomendadas según el número de casos con el que se trabaje.

2. El segundo interrogante que surge es, ¿qué hacer cuando hay diversidad limitada? **QCA** permite tres tipos distintos de soluciones dependiendo del tratamiento dado a los remanentes, que también se denominan

supuestos simplificadores (*simplifying assumptions*) **(Ragin, 2003b, p. 9).**

a. Una primera estrategia para el tratamiento de residuales, es obtener la solución compleja, por cierto, considerada por **Schneider y Wagemann (2012)** como la más conservadora ya que hace uso de la minimización de solo aquellas configuraciones que contengan información empírica, excluyendo así el uso de contrafactuales al llevar a cabo la *minimización Booleana* **(Schneider y Wagemann, 2007a, p. 107)**. La *solución compleja,* sería un subconjunto de otras posibles soluciones que, al ser obtenidas, respetarán la información empírica de los casos observados contenidos en la tabla **(Rihoux y De Meur, 2009; Schneider y Wagernann, 2012)**.

b. Una segunda estrategia es *seleccionar la solución más parsimoniosa* **(Schneider y Wagemann, 2007b, p. 106)**, puesto que en el proceso de minimización el software incluirá los residuales que permitirán la obtención de esta fórmula. En la *solución parsimoniosa* se utilizarán todos los contrafactuales para producir la minimización. Es decir, se utilizarán todas las *posibles combinaciones de condiciones*, hayan ocurrido en la realidad o no. Y además, las combinaciones de casos que no hayan ocurrido se utilizarán con independencia de si son teóricamente verosímiles o no. Por eso, como advirtieron **Ragin y Sonnett (2004, p. 14)**, mientras que el uso de esta solución favorece la parsimonia, puede conducir a resultados tan parsímoniosos que acaben siendo poco realistas debido a la incorporación de contrafácticos improbables. Además, la *solución parsimoniosa* implica otro riesgo: las potenciales *contradicciones en* el uso de los residuales, puesto que se puede incluir un mismo residual tanto para la explicación de la presencia del resultado como para su *ausencia* **(Rihoux y De Meur, 2009, p. 64)**. Si esto ocurriera, siempre se debe tener en cuenta que aun así la *solución parsimoniosa* jamás va a contradecir la información empírica contenida en la *tabla de verdad*. En cualquier caso, este par de limitaciones aconseja usar las *soluciones parsimoniosas* de forma muy prudente.

c. Una tercera estrategia, es obtener la solución intermedia **(Ragin, 2008, p. 147; Schneider y Wagemann, 2012)** mediante el uso de contrafactuales fáciles (*easy counterfactuals*) y de la indicación de *expectativas direccionales*. Las primeras son supuestos simplificadores que están en línea tanto con la evidencia empírica como con el conocimiento teórico existente sobre los efectos de las condiciones individuales que componen el remanente lógico **(Schneider y Wagemann, 2012, p. 168)**. Los segundos son argumentos justificados y derivados teóricamente de que una condición individual debiera contribuir a la ocurrencia de un resultado cuando este está presente en vez de ausente (o viceversa) **(Schneider y Wagemann, 2012, p. 325)**. A efectos prácticos, los residuales se incluyen en la minimización solo si hay razones teóricas o empíricas para creer que, aunque no ocurrieron casos empíricos con esas configuraciones contrafáctuales, si hubieran ocurrido hubieran conducido al *resultado de interés* **(Ragin, 2008)**. La *solución intermedia* permite por tanto abordar el problema de la *diversidad limitada* en el análisis configuracional, complementando las configuraciones observadas con configuraciones no observadas, pero solo en la medida en que sean verosímiles. Unas y otras serán incluidas en la *minimización*, favoreciendo la *parsimonia* de las soluciones aunque de manera controlada y guiada por el conocimiento teórico y empírico existente sobre el objeto de estudio. Aunque son diferentes, estas tres clases de soluciones subrayan diferentes dimensiones de los casos explorados, y tienen cada una de ellas lecturas distintas. Sin embargo, en los últimos tiempos se considera el uso de *la solución intermedia*, en general, como una buena práctica **(Ragin, 2009b, p. 111)**.

Para resumir cómo **QCA** enfrenta la *diversidad limitada*, se presenta a continuación, en la **Tabla 3.9**, el siguiente esquema de la lógica del proceso del Análisis Estándar (**AE**) **(Ragin y Sonnett, 2004; Ragin, 2008; Schneider y Wagemann, 2012)** para tratar residuales y obtener las distintas soluciones.

Tabla 3.9. La lógica del proceso de análisis estándar (AE)

Parámetros	*Solución compleja*	*Solución parsimoniosa*	*Solución intermedia*
Selección de los residuales	Los residuales lógicos se excluyen del análisis	Todos son lncluidos *a priori*, aunque el software se encargará de procesarlos	El investigador escoge algunos residuales para la minimización según decisiones justificadas
Propósito	Versosimilitud. Solo configuraciones con evidencia empírica entran al proceso de minimización	Parsimonia. Busca maximizar la reduccción de la complejidad	Verosimiltud teórica y parsimonia
Solución	Se obtiene solución más compleja posible	Se obtiene la solución más parsimoniosa posible	Se sitúa en un punto intermedio entre las soluciones más compleja y parsimoniosa

Fuente: Schneider y Wagemann (2012) con adaptación de Medina et al., (2017).

Cabe decir que el investigador puede reunir al Análisis Estándar Mejorado (**AEM**, por sus siglas en español), que surge como respuesta a dos fallos presentes en el **AE**, a saber:

1. La obtención de soluciones basadas en *supuestos insostenibles*, y
2. La selección única de *residuales* que contribuyen a la *parsimonia* dejando de lado la opción de evaluar los residuales por su importancia y contribución a la teoría **(Schneider y Wagemann, 2012, p. 197)**.

El **AEM** permite restringir la inclusión de *supuestos insostenibles* en las soluciones provistas por **QCA**. Las *supuestos insostenibles* pueden ser *supuestos incoherentes o implausibles*, es decir, *supuestos* que contradicen las afirmaciones hechas acerca del mismo remanente en diferentes momentos del proceso analítico **(Schneider y Wagemann, 2012, p. 198)**. Por ejemplo, si se incluye el mismo residual para explicar la presencia y la ausencia del resultado, o se hacen *supuestos* acerca de remanentes imposibles que contradicen el sentido común **(Schneider y Wagemann, 2012, p. 199)**.

Existen formas de evitar los *supuestos insostenibles*. Una estrategia trataría de excluir aquellos *residuales* que contribuyan a crearlas. Las ventajas de ello son que, por un lado, ninguna solución se basará en *supuestos insostenibles*, impidiendo que se formulen *supuestos simplificadores contradictorios* y, por otro lado, que las condiciones necesarias no desaparecerán **(Schneider y Wagemann, 2012, p. 209)**. Así, el **AEM** permite obtener la solución *más parsimoniosa mejorada* (*enhanced most parsimonious solution*), así como la *solución intermedia mejorada* (*enhanced intermediate solution*).

Como ejemplo, vamos a aplicar el **AEM** a los datos de la **Tabla 3.8.**

1. El primer paso es *excluir* cualquier *residual* del proceso de *minimización* que pudiera generar un *supuesto insostenible*. Para ello, vamos a realizar el *análisis de necesidad* de cada una de las condiciones (tanto para la presencia del resultado como para su ausencia) y compararlos.

2. Con ayuda del *software*, se detectó como *necesaria* para **(Y)** la condición **STL** con consistencia de **1.0** y cobertura del **66%** (**STL** ← **Y**).

- **a.** No se observan condiciones necesarias para **(~Y).** A partir de este análisis, se deben excluir del proceso de minimización de **(Y)** los dos residuales de la *tabla de verdad* hipotética, pues estos *contradicen* la afirmación de necesidad.
- **b.** Si hubiera habido alguna condición necesaria para **(~Y),** la configuración asociada a ella también hubiera tenido que suprimirse del análisis de **Y,** para *evitar contradicciones.*
- **c.** En cualquier caso, las expectativas direccionales que señalaremos para obtener la *solución intermedia*, siguiendo el resultado obtenido en el análisis de necesidad para **Y**, debe plantear que la condición **STL** debe estar presente para la ocurrencia del fenómeno.
- **d.** La misma lógica y procedimiento es aplicado al análisis de la ausencia del resultado **(~Y).** Se deben suprimir los residuales incluidos en el análisis del resultado **Y** para *no contradecir* las afirmaciones, así como aquellos que contengan los resultados de necesidad de **Y.**

e. Además, el investigador puede aplicar la Ley de De Morgan al resultado obtenido en el análisis de necesidad de **Y** para formular los supuestos en la obtención de la *solución intermedia*.

f. En cuanto a las soluciones, *la compleja y la parsimoniosa* son idénticas porque los dos residuales existentes fueron excluÍdos del análisis:

(STL * DSY) + (STL * BSY)

Nota: El cual cubre los casos: **A,B,C,D; A,B,C,M**

g. La solución intermedia (asumiendo la presencia de **STL**), se presenta de la siguiente manera**:**

(BSY * STL) + (DSY * STL)

Nota: El cual cubre los casos: **A, B, C, D; A, B, C, M**

Más allá del **AEM**, *existe otra estrategia* para el tratamiento de remanentes llamado *Análisis Estándar Mejorado Guiado por la Teoría* (**AEMeT o TESA** *Theory-Guided Enhanced Standard Analyses*). Esta estrategia, restringe el rango de residuales utilizados y trae de vuelta a aquellos supuestos insostenibles. El objetivo de esta práctica es *olvidarse de la premisa central de que la parsimonia es el principio rector para la elección de residuales para los contrafactuales* **(Schneider y Wagemann, 2012, p. 212)**.

A diferencia del **AME**, el **AEMeT** prioriza la *factibilidad teórica en la selección de los residuales* para suponer la existencia de contrafactuales que podrían (o no) *reducir la parsimonia*, aunque al igual que para el **AME**, el **AEMeT** solo hace uso de *supuestos defendibles* **(Schneider y Wagemann, 2012)**. En términos simples, el fundamento del **AEMeT** es introducir filas de la *tabla de verdad* como buenos *contrafactuales*, pero *no simplificadores*, en el proceso de *minimización*, es decir, no orientados a la *parsimonia*, para lo que se recomienda formular expectativas direccionales de conjunciones de condiciones en vez de condiciones individuales **(Schneider y Wagemann, 2012, p. 217)**.

Veamos otro ejemplo dado que desafortunadamente, el mundo empírico ofrece relativamente pocas oportunidades para construir comparaciones totalmente articuladas, como comparaciones basadas en experimentos. Incluso formas muy simples de análisis causal, se ven obstaculizadas por la diversidad limitada. Considere por ejemplo, la **Tabla 3.10.**

Tabla 3.10. Ejemplo hipotético del impacto de la diversidad limitada

Empresas innovadoras clúster manufactura	Columna			
	Fuertes prácticas de innovación por mercadotecnia	Fuertes prácticas de innovación por el modelo de negocios	Fuerte posicionamiento de mercado	Universo
	U	L	G	N
1	Si	Si	Si	4
2	Si	No	No	6
3	No	No	No	2
4	No	Si	¿?	0

Fuente: Elaboración propia.

Que muestra, un conjunto de datos de forma hipotética de las mismas empresas innovadoras clúster sobre dos *condiciones causales*:

a. Fuertes *prácticas de innovación por mercadotecnia* **U** (sí / no)
b. Fuertes *prácticas de innovación por el modelo de negocios* **L** (sí / no), y
c. Un resultado, fuerte *posicionamiento de mercado* **G** (sí / no).

La tabla presenta:

1. Las cuatro combinaciones de las dos *condiciones causales* de presencia / ausencia, pero solo existen tres de las cuatro.
2. Específicamente, ninguna empresa innovadora clúster existente combina la ausencia de práctica fuerte de *innovación por mercadotecnia* **U** con la presencia de práctica de *innovación por el modelo de negocios* **L,** que produzca un fuerte *posicionamiento de mercado* **G.**
3. Una inspección simple de la tabla, revela una correlación perfecta entre la presencia fuertes prácticas de *innovación por mercadotecnia* **U** y

presencia de fuertes prácticas de *innovación por el modelo de negocios* **L**, que produce un fuerte *posicionamiento de mercado* **G,** sugiriendo una simple explicación *parsimoniosa*.

4. Observe sin embargo, que un enfoque alternativo a la evidencia, produce una explicación diferente. Si la pregunta es, ¿cuáles son las *condiciones causales relevantes* compartidas por todas las instancias para obtener el resultado fuerte *posicionamiento de mercado* **G**?

5. La respuesta se encuentra en dos condiciones compartidas: la presencia de fuertes prácticas de *innovación por mercadotecnia* **U** y la presencia de fuertes prácticas de *innovación por el modelo de negocios* **L.**

6. Además, ninguno de los casos negativos (instancias con ausencia del fuerte *posicionamiento de mercado,* **G**) comparten esta combinación.

7. Esta segunda estrategia analítica, indica que la presencia de fuertes prácticas de *innovación por mercadotecnia* **U** y la presencia de fuertes prácticas de innovación por el modelo de negocios **L,** son las que explican el surgimiento del fuerte *posicionamiento de mercado* **G** y no las fuertes prácticas de *innovación por el modelo de negocios* **L** por sí misma.

8. ¿Cuál explicación es correcta? Un análisis cuantitativo convencional de estos datos, apuntan a la primera explicación porque no es solo más *parsimoniosa*, sino que también, está completo a partir del punto de vista de la varianza de que no hay casos inexplicables.

9. Sin embargo, los investigadores orientados a casos, no son tan afectos a explicaciones basadas en el uso de la *parsimonia*, sino más bien, prefieren las explicaciones por causalidad abocadas con lo que saben ellos mismos de los casos.

10. Por lo general, cuando los casos se examinan en profundidad, los investigadores encuentran que la causalidad es compleja y muy a menudo implica, combinaciones de *condiciones causales* (o trayectorias causales).

11. Sin duda, favorecerían la segunda explicación sobre la primera por los investigadores orientados a casos, con bases analíticas. La búsqueda de elementos comunes causalmente relevantes compartidos por un

conjunto de casos con el mismo resultado, es frecuentemente, el primer movimiento analítico en la investigación orientada a casos, a pesar del hecho de que esta práctica de seleccionar en la variable dependiente, es casi universalmente condenado por investigadores cuantitativos que piensan solo en términos de correlaciones.

12. En un nivel más formal, cuál respuesta es correcta, depende del resultado que se observaría para los casos que exhiban la ausencia de fuertes prácticas de *innovación por mercadotecnia* **U** combinada con la presencia de fuertes prácticas de *innovación por el modelo de negocios* **L**, es decir, si se pudieran encontrar tales casos.

13. Si estos casos muestran un fuerte *posicionamiento de mercado* **G,** entonces la conclusión sería que tener presencia de fuertes prácticas de *innovación por el modelo de negocios* **L,** por sí mismo, generan un fuerte *posicionamiento de mercado* **G.**

14. Si estos casos muestran falla al mostrar un fuerte *posicionamiento de mercado* **G,** entonces la conclusión sería que es la combinación de presencia de fuertes prácticas de *innovación por mercadotecnia* **U** con la presencia de fuertes prácticas de *innovación por el modelo de negocios* **L**, las que explican un fuerte *posicionamiento de mercado* **G.**

15. Si son casos relevantes que combinan presencia de fuertes prácticas de *innovación del modelo de negocios* **L** pero no presencia de fuertes prácticas por *innovación por mercadotecnia* **U** que no pudieron ser identificadas, entonces los investigadores deben preguntarse: ¿qué pasaría en tales casos?, ¿podría darse un fuerte posicionamiento de internet?

16. Para responder estas preguntas, los investigadores deben confiar en su conocimiento teórico y sustantivo, que a su vez proporcionaría la base para decidir entre las dos explicaciones:

a. La *parsimoniosa* (*causa única*) vs

b. Las más complejas (*causas combinadas*)

17. En resumen, la elección de las explicaciones es dependiente de la teoría y conocimiento empírico y sustantivo.

18. Tenga en cuenta que aunque el ejemplo es muy simple, hay solo dos *condiciones causales,* y solo una de las cuatro combinaciones causales

carece de casos; por lo que es imposible sacar una conclusión firme sobre causalidad directamente de la evidencia presentada, debido a la limitada diversidad de casos empíricos.

19. Además, qué respuesta es correcta, a los ojos de la ciencia social contemporánea, podría ser una cuestión de gusto. Los estudiosos que favorecen la *parsimonia* pueden preferir la primera respuesta; los académicos que buscan una conexión más cercana a los casos podrían preferir el segundo. Es importante señalar que esta ambigüedad con respecto a cuál es la explicación correcta, no es una consecuencia del uso de *condiciones causales* dicotómicas.

Casos contrafactuales

Evaluar un resultado factible, de una combinación de *condiciones causales* que no existen, en el que su origen es la imaginación del investigador, puede ser muy difícil de entender. Sin embargo, esta estrategia analítica tiene una larga y muy importante tradición en el historia de las ciencias sociales. Una combinación causal, que carece de instancias empíricas y por lo tanto, es producto de la experiencia e imaginación del investigador, es un *caso contrafactual*; evaluar su resultado factible, es el *análisis contrafactual* **(Hicks et al., 1995)**.

Para algunos, el *análisis contrafactual* es fundamental para la *investigación orientada a casos*, porque dicha investigación generalmente abarca solo un puñado empírico de casos **(Fearon, 1991)**. Si solo existieran unos pocos casos (por ejemplo, de innovación abierta) entonces los investigadores deben comparar los casos empíricos vs. los casos hipotéticos. La afinidad entre el *análisis contrafactual* y la investigación orientada a caso, sin embargo, se deriva no solo de su enfoque en **N** pequeñas muestras, sino de su naturaleza *configuracional*. Las explicaciones de los resultados de investigaciones orientadas a casos, a menudo son de naturaleza combinatoria, enfatizando configuraciones específicas de condiciones causales.

Los casos *contrafactuales*, por lo tanto, a menudo difieren de los casos empíricos por una sola *condición causal*, creando así una comparación

decisiva, aunque parcialmente imaginaria. La consideración de casos *contrafactuales*, es a menudo explícita en la investigación comparativa orientada a casos en casos hipotéticos, dada la ventaja de usar su capacidad de combinar, tanto la retórica como la analítica en narrativas útiles de casos, como los de innovación.

Sobre argumentar casos hipotético, **Max Weber (1949)** es comúnmente citado como el primer científico en abogar por el uso de *experimentos del pensamiento*, generando supuestos en la investigación social. Argumentó que los investigadores pueden obtener información sobre la importancia causal de componentes individuales, de eventos mediante la realización de experimentos del pensamiento, que imagina casos irreales. La visión de **Weber (1949, p. 187),** se basa en una configuración explícita del enfoque del análisis causal, en que:

> *"...un resultado concreto, no puede ser visto como el producto de una lucha de ciertas causas, que la favorecen y otras que se le oponen. La situación debe, en cambio, verse como la totalidad de todas las condiciones que van a la cadena causal, desde el efecto y que habrían actuado conjuntamente, de cierta manera y de ninguna otra manera, para que se realice el efecto concreto..."*.

Los investigadores comparativos contemporáneos, han seguido debatiendo sobre cómo construir y usar los *contrafactuales* en la investigación y en el desarrollo de teorías **(Fearon, 1991; Hawthorn, 1991; Tetlock y Belkin 1996a)**. En la introducción de su volumen sobre experimentos de *pensamiento contrafactual*, **Tetlock y Belkin, (1996b, p. 4)** describen cinco estilos de *argumentación contrafactual* y sugieren seis criterios que utilizan los investigadores para juzgar estos argumentos. Aunque los estilos descritos *contrafactuales* tienen un rango amplio de argumentos, ninguno formaliza el uso de *contrafactuales*, dentro de una comprensión explícitamente *configuracional* de la causalidad.

Como se observa a continuación, el marco de configuración del análisis cualitativo comparativo (**QCA.** *Qualitative Comparative Analysis*) ofrece una guía útil para usar *contrafactuales* en la investigación social con orientación a la administración de la innovación. En **QCA**, los casos *contrafactuales* se conciben *como sustitutos de los casos empíricos coincidentes*. Estos casos hipotéticos coincidentes, se identifican por sus *configuraciones* de condiciones de causalidad. En un nivel más abstracto, el análisis contrafactual está implicado cada vez que un investigador, hace una inferencia causal basada en el análisis de datos sociales naturales (es decir, *no experimentales*); datos cuya diversidad limitada es la norma.

Por ejemplo, de la **Tabla 3.10**, cuando los investigadores afirman que la fuerte práctica de la *innovación por el modelo de negocios* **L** es una causa importante de *fuerte posicionamiento de mercado* **G** neto, de otras causas relevantes, argumentando en efecto, que las empresas innovadoras clúster con débiles o nulas prácticas de *innovación por el modelo de negocios* **L** (por ejemplo, en clúster de países emergentes), tendrían más *fuerte posicionamiento de mercado* **G**, si solo esta característica fuera diferente.

Así, la interpretación del efecto observado, invoca a las empresas innovadoras clúster hipotéticas, por ejemplo, una empresa hipotética que figure en alta puntuación de *innovación pero de alta tecnología software* **S** (no manufacturera como es el caso) fuera del clúster, en todos los aspectos causalmente relevantes, excepto en la que tiene una fuerte práctica de *innovación por el modelo de negocios* **L.** Obviamente, como entidades no experimentadoras, los científicos sociales no pueden crear dicha empresa, por lo que no pueden asignar *condiciones causales* a sus casos como una entidad experimentadora, distribuiría sus tratamientos, entre sujetos aleatorizados. Ellos están atrapados con datos no experimentales y deben contender con el hecho de que, una variedad de factores observados y no observados usualmente participan en procesos de selección natural (por ejemplo, la creación de la empresa innovadora hipotética, para introducirla en el clúster). Estos procesos de selección natural, a su vez, distorsionan la estimación de los efectos causales (por ejemplo, el impacto de la fuerza de la empresa inno-

vadora hipotética en sus prácticas de *innovación por el modelo de negocios* **L,** en el *fuerte posicionamiento del mercado*, **G**.

El problema de la selección, ha llevado a econométricos y estadísticos a desarrollar un marco general para comprender la causalidad, en términos de la diferencia entre el valor de cada caso sobre la variable dependiente, cuando está en las condiciones de control vs. las condiciones de tratamiento (por ejemplo, la empresa innovadora hipotética clúster con y sin fuertes prácticas de *innovación por el modelo de negocios*, **L**). Por otro lado, **Winship y Morgan (1999, p. 660)** sostienen que el lenguaje del tratamiento y control de las variables, son generalmente aplicables: en casi cualquier situación, en la que un investigador intenta estimar un efecto causal, el análisis puede describirse, al menos en términos de un experimento del pensamiento, como un experimento. Una implicación más directa del uso del lenguaje experimental, que no se discute en detalle, es la restricción de que el tratamiento debe ser manipulable.

Holland (1986), argumenta que no tiene sentido, hablar sobre el efecto causal del género o cualquier otro rasgo individual, que no sea manipulable por sí mismo. Uno debe modelar, explícitamente el mecanismo manipulable que genera un efecto causal aparente, de un atributo no manipulable **(Winship y Morgan, 1999, p. 663)**.

Dado que solo una de estas dos condiciones es observable, la otra debe estimarse estadísticamente, teniendo en cuenta los efectos de los procesos de selección **(Holland, 1986; Sobel 1995; Winship y Morgan, 1999; Winship y Sobel, 2004; Brady 2003)**. Aunque sofisticados en su enfoque del problema de la selección, los procedimientos de *regresión contrafactual* son factibles solo cuando:

1. Hay una **N** de muestras muy grande, y
2. Es factible *a priori*, que cada caso, pudiera estar o en el grupo de control o en el grupo de tratamiento **(Winship y Morgan, 1999)**.

Además, estos procedimientos, como el análisis estadístico convencional, permanecen lineales y aditivos, por lo que no se examinan los

problemas de *diversidad limitada* y casos coincidentes directamente. Un intento de análisis de la diversidad limitada o la llamada maldición de la dimensionalidad, con *regresión Booleana probit y logit*, es tratada en **Braumoeller (2003)**.

El problema de los casos coincidentes

Para apoyar un argumento, que enfatiza combinaciones de *condiciones causales*, los investigadores deben comparar los casos que coinciden estrechamente uno con el otro. La comparación ideal, siguiendo a **Mill (1967)**, es entre pares de casos que difieren en una sola *condición causal*. Tales comparaciones, ayudan a los investigadores a establecer si una *condición causal* específica, es una parte integral de la combinación de *condiciones causales*, que genera el resultado en cuestión. Es muy difícil hacer coincidir casos empíricos de esta manera, debido a la *diversidad limitada* de fenómenos empíricos. Por ejemplo, de la misma **Tabla 3.10** para interpretar el impacto de tener fuertes prácticas *por innovación del modelo de negocios* **L,** con resultado a un *fuerte posicionamiento de mercado* **G.** Lo ideal, es realizar una comparación con una empresa innovadora del mismo clúster y verificar las causas del porqué se genera dicho *fuerte posicionamiento de mercado* **G** pero con una *empresa innovadora de alta tecnología de software* **S,** fuera del clúster. La búsqueda de casos coincidentes, depende profundamente de lo que sustente el investigador, porque el proceso de emparejamiento o coincidencia, debe centrarse en las *condiciones causales* identificadas según sus conocimientos teóricos y conocimientos empíricos sustantivos.

Para ilustrar el papel de los casos empíricos coincidentes, considere una investigación *orientada a caso*, donde el investigador argumenta, que en empresas altamente innovadoras, cuatro *condiciones causales* se combinan para producir un *fuerte posicionamiento de mer*cado, **G**:

a. Fuertes prácticas de *innovación por mercadotecnia,* **U.**

b. Fuertes prácticas de *innovación por el modelo de negocios,* **L**.
c. Robusta *infraestructura de soporte de internet* **H**.
d. Alta *visión del CEO y gerencias sobre uso tecnologías de información* **C**.

El investigador cita a los cinco empresas más avanzadas tecnológicamente en el mundo, como instancias relevantes de este argumento. Este argumento causal, requiere (al menos) de cuatro tipos de casos muy parecidos, por ejemplo: empresas altamente innovadoras similares a las cinco empresas más avanzadas tecnológicamente en el mundo, pero *sin robusta infraestructura de soporte de internet* **H,** o empresas altamente innovadoras similares a las cinco empresas más avanzadas tecnológicamente en el mundo, pero sin alta *visión del CEO* y *gerencias sobre uso tecnologías de información* **C** y así. Los casos de comparación coinciden con las cinco empresas más avanzadas tecnológicamente en el mundo en tres de las cuatro *condiciones causales*. Estos casos coincidentes, se pueden representar usando álgebra booleana, de la siguiente manera:

1. Los casos de las cinco empresas más avanzadas tecnológicamente en el mundo:

$$\mathbf{(H * C * L * U) <= G}$$

2. Los cuatro casos coincidentes:

$$\mathbf{(\sim H * C * L * U) + (H * \sim C * L * U) + (H * C * \sim L * U) + (H * C * L * \sim U) <= \sim G}$$

Donde además:

a. ~ indica ausencia o negación;
b. (*) indica combinación de condiciones (lógica **AND**),
c. (+) indica combinaciones alternativas de condiciones (lógica **OR**),
d. ($\leq$) indica es suficiente para...

e. Si el investigador es capaz de demostrar que un fuerte *posicionamiento de mercado* **G**, no se logró desarrollar en los cuatro casos coincidentes, este hallazgo reforzaría en gran medida su argumento causal. En efecto, la ausencia del resultado en estos cuatro casos coincidentes, permitirían al investigador afirmar que cada uno de las cuatro *condiciones causales*, es una condición **INUS** (*lnsufficient but Necessary part of a condition which is itself Unnecessary but Sufficient for the result*) **(Mackie 1965, 245)**.

f. Los casos combinados ideales son, a menudo, difíciles de encontrar para algunas combinaciones de *condiciones causales* que son poco probables y otras pueden ser empíricamente imposibles. Por ejemplo, puede resultar muy difícil identificar un una empresa innovadora clúster con:
1. Fuertes prácticas de *innovación por mercadotecnia,* **U.**
2. Fuertes prácticas de *innovación por el modelo de negocios,* **L.**
3. Robusta *infraestructura de soporte de internet,* **H.**
Pero nula *visión del CEO y gerencias sobre uso tecnologías de información,* **C.**

g. Además, cuando los argumentos causales son combinatoriamente complejos (es decir, que hay un resultado común, cuando los investigadores examinan casos en profundidad), el arreglo de la matriz de casos coincidentes, necesario para apoyar un argumento causal, puede ser sustancial.

h. El mundo empírico actual, está profundamente limitado en su diversidad, y los casos que coinciden en todas las *condiciones causales relevantes*, salvo una, son relativamente raras. Así, los investigadores comparativos, por lo general no pueden identificar los casos empíricos relevantes coincidentes y por lo tanto, deben sustituir por casos *contrafactuales*.

Casos contrafactuales vs. QCA

El análisis cualitativo comparativo (**QCA**), es una de las pocas técnicas disponibles hoy, que aborda directamente la diversidad limitada de la ocurrencia natural del fenómeno social. A diferencia de las técnicas convencionales, **QCA** comienza:

1. Suponiendo que la causalidad sea *configuracionalmente compleja*, en lugar de ser simple. La mayoría de las técnicas convencionales suponen que las *condiciones causales* son variables independientes cuyos efectos sobre los resultados, son lineales y aditivos. La clave de **QCA** es que ve los casos como *configuraciones* de *condiciones causales* y utiliza *tablas de verdad* para representarlas y analizarlas lógicamente.

2. La **Tabla 3.10** es, de hecho, una *tabla de verdad* muy simple con dos *condiciones causales* y cuatro combinaciones causales.

3. En el lenguaje de **QCA**, la cuarta fila de la *tabla de verdad* que se muestra en la **Tabla 3.10** es un *residual* (una combinación de *condiciones causales* que carecen de instancias empíricas). En **QCA**, la solución a esta *tabla de verdad* depende sobre cómo se trata el *residual*.

4. La estrategia más conservadora es tratarlo como un falso (excluido) al evaluar las *condiciones causales*, para la aparición de una fuerte posicionamiento en el mercado **G** y también como falso (excluido), al evaluar las condiciones para la ausencia de fuerte posicionamiento en el mercado **G**, como sigue:

- Presencia de fuerte *posicionamiento en el mercado*:

$$(L * U) <= G$$

- Ausencia de fuerte *posicionamiento en el mercado*:

$$(\sim L * U) + (\sim L * \sim U) <= \sim G$$

$$(\sim L) * (U + \sim U) <= \sim G$$

$$(\sim L) <= (\sim G)$$

Donde, las notaciones son las mismas que en el ejemplo anterior. La primer declaración resume la primera fila de la **Tabla 3.10**; la segunda declaración resume la segunda y tercera fila; las declaraciones tercera y cuarta, simplifican el segundo, usando álgebra booleana.

5. De acuerdo con este análisis, la combinación de fuertes prácticas de *innovación por el modelo de negocios* **L** y fuertes prácticas de *innovación por mercadotecnia* **U,** es suficiente para la aparición de fuerte *posicionamiento de mercado,* **G**.

6. La ausencia de fuertes prácticas de *innovación por el modelo de negocios* **L,** es suficiente para la ausencia fuerte *posicionamiento de mercado,* **G**.

En **QCA**, una estrategia alternativa, es tratar los residuos como si no les importara las combinaciones causales. La etiqueta neutral (*Don´t care*), refleja el origen del enfoque de la *tabla de verdad*, proveniente del diseño y análisis de circuitos de conmutación. Cuando es tratado el estado neutral (*Don't care*), el *residual* está disponible como un potencial simplificador de suposición. Es decir, se tratará como una instancia del resultado, si al hacerlo da como resultado una solución lógicamente más simple.

Del mismo modo, también puede tratarse como una instancia de la ausencia del resultado, nuevamente, si hacerlo da como resultado una solución lógicamente más simple para la ausencia del resultado, como sigue:

1. El uso, de *neutral* puede representarse en forma simbólica con el término *residual*:

$$(L * \sim U)$$

agregado ambas declaraciones:

- Presencia de fuerte *posicionamiento de mercado* **G**.

$$(L * U) + (L * \sim U) <= (G)$$
$$(L) * (U + \sim U) <= (G)$$
$$(L) <= (G)$$

- Ausencia de fuerte *posicionamiento de mercado* **~ G**.

$$(\sim L * U) + (\sim L * \sim U) + (L * \sim U) <= (\sim G)$$
$$(\sim L) * (U + \sim U) + \sim U * (L + \sim L) <= (\sim G)$$
$$(\sim L) + (\sim U) <= (\sim G)$$

2. A partir de estos resultados, queda claro que el uso de *residuales* como la combinación de neutrales, en la solución para la presencia de fuerte *posicionamiento en el mercado* **G,** conduce a una solución lógicamente más simple, mientras que conduce a una solución más compleja, para la ausencia de fuerte *posicionamiento en el mercado,* (**~ G**).
3. Por lo tanto, un investigador interesado en obtener una solución más *parsimoniosa*, podría preferir el uso de los *residuales* (la cuarta fila de la *tabla de verdad*) como una combinación de las neutrales, en solución para la presencia de fuerte *posicionamiento en el mercado,* **G**.
4. Tenga en cuenta, que el uso de los *residuales* como una combinación de los neutrales en el análisis de la presencia de un fuerte *posicionamiento en el mercado* **G,** ofrece el mismo resultado como un análisis estadístico convencional de los mismos datos.
5. Usando **QCA**, corresponde al investigador evaluar la factibilidad de cualquier combinación de neutrales que se incorpore en una solución. Suponga que el investigador en este ejemplo eligió la solución más *parsimoniosa* para la presencia de fuerte *posicionamiento en el mercado,* **G.**
6. Así, se concluye que este resultado, se debe enteramente a la presencia de fuertes prácticas de *innovación por el modelo de negocios,* **L**. Entonces es necesario que el investigador, evalúe la factibilidad del supuesto simplificador que esta solución incorpore, a saber, sobre si los casos de la presencia de fuertes prácticas de *innovación por el modelo de negocios* **L,** combinado con la ausencia de fuertes prácticas de *innovación por mercadotecnia* **U**, hicieran de hecho, que existan los casos que se mostraran con fuerte *posicionamiento en el mercado,* **G**.
7. Esta es una suposición muy fuerte. Muchos investigadores lo encontrarían inverosímil a la luz del conocimiento empírico sustantivo y de marco teórico. Ese conocimiento existente en parte, sería el simple hecho de que todos los casos conocidos de fuerte *posicionamiento en el mercado* **G**, (en este ejemplo hipotético) ocurren en contextos de empresas con fuertes prácticas de *innovación por el modelo de negocios,* **L.** El conocimiento existente también incluiría el análisis en

profundidad, a nivel de caso la aparición fuerte *posicionamiento en el mercado,* **G.** Este conocimiento podría indicar, por ejemplo, que se han establecido centralmente fuertes prácticas de *innovación por mercadotecnia* **U,** involucrado en el establecimiento de fuerte *posicionamiento en el mercado,* **G.**

8. El punto importante aquí, no es la conclusión específica del estudio o si tener o no fuertes prácticas de *innovación por el modelo de negocios* **L** es suficiente por sí solo para establecimiento de fuerte *posicionamiento en el mercado,* **G.** Más bien, el problema es el estado de suposiciones acerca de las combinaciones de *condiciones causales,* que carecen de instancias empíricas. En **QCA**, estos supuestos deben ser evaluados; las combinaciones *neutrales* (*residuales*) no deben incluirse en soluciones a modo *mecanicista* porque, después de todo, sí importan.

Contraste con la investigación convencional cuantitativa

En la investigación cuantitativa convencional, el tema de la *diversidad limitada* no suele ser tomado en cuenta, porque los investigadores utilizan técnicas y modelos en los que adoptan supuestos muy específicos, sobre la naturaleza de la causalidad. Típicamente, los investigadores comienzan su investigación desarrollando listas de posibles causas de factores relevantes, para la obtención del resultado en cuestión, utilizando una variedad de perspectivas teóricas. Generalmente, tratan cada *condición causal* como una causa independiente del resultado y consideran como su principal tarea analítica, evaluar cuál de las *condiciones causales* enumeradas son los más importantes. Es decir, intentan identificar los mejores predictores del resultado, basados en estimaciones estadísticas de efecto neto de cada variable. La estimación de los *efectos netos,* a su vez, está basado sobre el supuesto de que cada causa, por sí sola, es capaz de influir el resultado; es decir, supone que las causas son independientes y aditivas en sus efectos. Por lo tanto, la investigación cuantitativa convencional evita el problema de la diversidad limitada al suponer que la causalidad es irrealmente simple.

QCA, por el contrario, hace énfasis al carácter combinatorio de la *investigación orientada a casos*, es decir, a la idea de que la causalidad puede ser compleja y que el mismo resultado puede obtenerse de diferentes combinaciones de las *condiciones causales*. Esta idea es implementada a través de elaborar *tablas de verdad*, que consideran todas las combinaciones lógicamente posibles de las *condiciones causales relevantes*. Cuando el número de casos es pequeño a moderado, es común diseñar *tablas de verdad* con solo dieciséis filas (basada en cuatro *condiciones causales*), tener filas sin casos (es decir, filas *residuales*), que son todos los posibles casos *contrafactuales*. Sin embargo, tener una gran cantidad de casos, no es garantía de que se puedan evitar *los residuales.* De nuevo, la *diversidad limitada* (que se manifiesta con una abundancia de filas *residuales*) es la regla, no la excepción, en el estudio de los fenómenos sociales. Hay evidencias interesantes de casos que cumplen con la investigación, por ejemplo:

1. **Ragin (2003)**, demuestra que en un conjunto de datos grande (**N** = **758**) solo llena **24** filas de treinta y **2** en la *tabla de verdad* con **5** *condiciones causales* y que trece de estas **32** filas contienen, casi todos los casos **(96.7%)**.
2. **Sonnett (2004)**, en un análisis de datos a nivel individual sobre gustos musicales (**N** =**1606**), también encuentra que veintidós de sesenta y cuatro filas en la *tabla de verdad* (**34%** de las filas) contienen la mayor parte de encuestados en la muestra **(90%).**
3. **Braumoeller (2003, p. 229),** también encuentra evidencia de covarianza compleja (que denominó *diversidad limitada*) en un conjunto de datos con **8328** observaciones.

Desde este punto de vista, es fácil observar por qué el *análisis contrafactual* es esencial para la investigación en la innovación y en lo social. Cualquier análisis que permita la *complejidad combinatoria*, seguramente enfrentará una gran cantidad de *residuales* y, por lo tanto, un riqueza potencial de casos contrafactuales. La pregunta clave ha realizar, es ¿qué hacer al respecto? Se tienen las siguientes opciones:

a. Una opción, es retirarse al laboratorio y evitar datos no experimentales por completo. Este camino busca crear casos coincidentes a través de la manipulación experimental. La ruta del laboratorio conlleva severas restricciones en el tipo de preguntas que los científicos de innovación y sociales pueden hacer.

b. Otra opción, es utilizar técnicas estadísticas como las discutidas por **Winship y Morgan (1999)**, para estimar datos desconocidos (es decir, el valor ya sea de la condición de control o ya sea la condición de tratamiento), basado en un modelo estadístico que intenta controlar los procesos de selección subyacentes. La ruta estadística requiere no solo de un gran número de casos y un tipo específico de causalidad variable (una condición que puede ser manipulada), sino también, un número de suposiciones fuertes y simplificadoras sobre la naturaleza de la causalidad.

c. La tercera opción, es participar en un análisis contrafactual (por ejemplo, basado en experimentos del pensamiento). La ruta del experimento del pensamiento puede parecer poco atractiva porque implica lidiar con casos hipotéticos o poco factibles incluso irreales. En los capítulos por venir, sin embargo, se demostrará que muchos casos *contrafactuales*, pueden considerarse accesibles siempre que los investigadores tengan un sólido conocimiento teórico y empírico sustantivo bien desarrollado.

Suponga otro caso. Un investigador que postula, basado en un marco teórico y conocimiento empírico sustantivo, que los factores **(A, B, C y D)** son todas relevantes para el resultado **Y**, así:

1. La evidencia disponible indica, que muchas instancias de **Y** están acoplados con la presencia de *condiciones causales* **(A, B y C),** junto con la ausencia de la condición **D** (es decir, **A * B * C * ~ D <= Y**). Puede haber otras combinaciones no especificadas de *condiciones causales* vinculadas a un resultado **Y** en este ejemplo. No se supone, que esta sea la única combinación vinculado al resultado **Y.**

2. El investigador sospecha, sin embargo, que todo lo que realmente importa es la presencia de las tres primeras causas, **(A, B y C)** y que la cuarta condición **(~ D)** es superfluo en la presencia de **(A * B * C).** Sin embargo, no hay instancias de **(A, B y C)** combinadas con la presencia de **D** (es decir, no hay instancias de **A * B * C * D**). Por lo tanto, el caso coincidente decisivo para determinar si la ausencia o no de **D** es una parte esencial de la mezcla causal, simplemente no existe.
3. A través del *análisis contrafactual* (es decir, un *experimento del pensamiento*), el investigador podría declarar una combinación hipotética **(A * B * C * D)** para ser una instancia probable del resultado. Es decir, el investigador podría afirmar que **(A * B * C * D),** si existiera, conduciría a **Y**.
4. Este análisis contrafactual, permitiría la siguiente simplificación lógica:

(A * B * C * ~ D) + (A * B * C * D) <= (Y)

(A * B * C) * (~ D + D) <= (Y)

(A * B * C) <= (Y)

5. ¿Cuán factible es esta simplificación? La respuesta a esta pregunta depende del estado de conocimiento del marco teórico y empírico sustantivo relevante, sobre la conexión entre **D** e **Y** en la presencia combinada de las otras tres *condiciones causales* **(A * B * C).** Si el investigador puede establecer con la base del conocimiento existente, que existen razones para esperar que la presencia de **D** contribuya al resultado **Y** bajo estas condiciones (o viceversa, que la ausencia de **D** no debería de ser un factor contribuyente), entonces el análisis contrafactual presentado arriba es a la vez factible y razonable. En otras palabras, el conocimiento existente hace de la afirmación **(A * B * C * D) <= Y** se considere un contrafactual fácil, porque el investigador estaría agregando una contribución redundante de causa (la presencia de **D**) a una configuración que ya está pensada para conducir al resultado **(A * B * C).**
6. Es importante señalar, que lo que se ha logrado en este ejemplo (usando álgebra booleana) es una rutina, aunque a menudo implícita, en mucha investigación orientada a casos. Si los investigadores conven-

cionales orientados a casos, deben examinar la instancia **empírica (A * B * C * ~ D) <= (Y)**, lo harían probablemente desarrollando su argumento causal o narrativa en función de factores pensados que estén relacionado con el resultado (es decir, la presencia de **A, B y C**). Igualmente, bien podrían considerar la posibilidad de que la ausencia de **D** (es decir, **~ D**) observado en estos casos podría ser integral en alguno forma, a la producción de **Y** por **(A * B * C).**

7. Sería muy probable también, concluir lo contrario, dado el supuesto estado de conocimiento existente sobre las cuatro *condiciones causales*, a saber, que es la presencia de estos factores causales, no su ausencia, y que estén vinculados a la aparición del resultado. Por lo tanto, llegarían rápidamente a la conclusión más *parsimoniosa*, que (**A * B * C)<= Y**.

8. El punto aquí ,es que el análisis contrafactual no siempre es explícito o incluso, articulado en la investigación orientada a casos, especialmente cuando los *contrafactuales* son tan fáciles. Tales análisis, son conducidos rutinariamente por investigadores orientados a casos sobre la marcha, es decir, en el proceso de construir explicaciones de un caso o una categoría específica de casos.

Ahora considere la situación opuesta:

1. El investigador observa instancias de (**A * B * C * D) <= Y** pero cree que **D** es superfluo o redundante en la producción del resultado **(Y)** dada la presencia de **(A * B * C).** Así ¿qué pasaría si **D** estuviera ausente? Desafortunadamente, en este caso, no hay casos de **(A * B * C * ~ D),** y el investigador debe recurrir al análisis contrafactual.

2. Sin embargo, suponga que el conocimiento de marco teórico y empírico sustantivo existente, conecta la presencia de **D** con el resultado **Y** entonces ¿es razonable afirmar que **(A * B * C * ~ D),** si existiera, conduciría al resultado de **Y**? Este contrafactual, que permitiría la simplificación de **(A * B * C * D)** a **(A * B * C),** se considera un *contrafactual difícil*.

3. De esta forma, el investigador tendría que realizar un esfuerzo concertado, basado tanto en su conocimiento del marco teórico, para soportar con argumentación detallada así como de su conocimiento empírico sustantivo, que soporte el caso Las discusiones metodológicas de los *contrafactuales*, a menudo asumen una variante no configuración de la forma difícil (Fearon, 1991, p. 39):

> *"...al intentar discutir o evaluar si algún factor* ***A*** *causó el evento* ***B,*** *los científicos sociales usan con frecuencia contrafactuales. Es decir, preguntan si o afirman que si* ***A*** *no hubiera ocurrido, entonces* ***B*** *no habría ocurrido..."*

4. El punto aquí, no es que los *casos contrafactuales difíciles* deban evitarse; más bien, el punto es que requieren una explicación y justificación cuidadosas.
5. A veces los investigadores logran justificar sus *contrafactuales difíciles*, y por lo tanto, tales esfuerzos han conducido a importantes avances y conocimientos teóricos.
6. La *distinción contrafactual fácil vs. contrafactual difícil*, no es una dicotomía rígida, sino más bien un continuo de factibilidad. Es decir:
   - **a.** En un extremo hay *contrafactuales más fáciles*, que asumen que agregar una *condición causal* redundante a una configuración que se sabe produce el resultado (por ejemplo, la condición **D** para la combinación (**A * B * C**) aún produciría el mismo resultado.
   - **b.** En el otro extremo, se encuentran los *contrafactuales más difíciles*, que intentan eliminar la contribución de la *condición causal* de una configuración que produce el resultado, en el supuesto de que esta causa es redundante y la configuración reducida, aún produciría el mismo resultado.

   La ubicación exacta, de cualquier uso específico de un *contrafactual* en el continuo *fácil / difícil*, depende principalmente del estado del conocimiento del marco teórico y conocimiento de empírico sustantivo existentes en la comunidad científica tanto de innovación, como social en general.

7. Este conocimiento, ayuda al investigador a decidir qué causas pueden ser *redundantes o irrelevantes,* dando apoyo teórico o empírico a los *argumentos contrafactuales* sobre la importancia o irrelevancia de una *condición causal* particular **(Tetlock y Belkin, 1996b)**.
8. Este aspecto del *análisis contrafactual,* también destaca la dependencia de la teoría y el conocimiento de la investigación científica social particularmernte enfocada a la innovación en general, así como su naturaleza fundamentalmente colectiva y social **(Merton, 1973)**.
9. Debido a que, la *diversidad limitada* es la regla y no la excepción en el estudio de los fenómenos sociales naturales, habrá muchas combinaciones lógicamente posibles de *condiciones causales* que carezcan de instancias empíricas en la mayoría de las investigaciones científicas sociales.
10. Estos casos *contrafactuales*, se pueden utilizar para simplificar los resultados, como se acaba de demostrar. Algunos de estos *contrafactuales* serán relativamente *fáciles* (y, por lo tanto, más o menos rutinarios); algunos serán *difíciles* (y, por lo tanto, tal vez deberían evitarse).
11. La consideración clave es el soporte de conocimiento de marco teórico y empírico sustantivo, del investigador, que está subyacente a cada uso.

Los contrafactuales fáciles y difíciles

Los investigadores que utilizan el análisis cualitativo comparativo (**QCA**. *Qualitative Comparative Analysis*) tienen dos opciones principales cuando se enfrentan a la *diversidad limitada* y abundancia de *residuales* (y, por lo tanto, muchos posibles casos *contrafactuales*):

1. Pueden evitar usar cualquier *residual*, para simplificar una *tabla de verdad*; por lo que esta opción excluye los casos *contrafactuales* por completo; o
2. Pueden permitir la incorporación de un subconjunto de *residuales* que produzcan la solución más *parsimoniosa* de la *tabla de verdad*; lo que permite la inclusión casos *contrafactuales fáciles y difíciles*, sin ninguna evaluación de su factibilidad.

A primera vista, ninguna de estas opciones parece atractiva. Es probable que el primero, conduzca a resultados innecesariamente complejos; como:

1. En el primer ejemplo descrito anteriormente, donde la simple suma de un *contrafactual* fácil, permitiría la simplificación de

(A * B * C * ~ D) a (A * B * C)

2. El segundo, puede conducir a resultados *parsimoniosos* poco realistas debido a la incorporación de *contrafactuales difíciles,* como en el segundo ejemplo, donde (**A * B * C * D**) se simplifica a (**A * B * C**).

Sin embargo, en lugar de rechazar estas dos opciones, es importante considerarlos como los puntos finales de un *continuum* de posibles resultados. Un extremo del *continuum* apunta a la *complejidad* (ningún caso *contrafactual* está permitido); el otro extremo apunta a la *parsimonia* (todos los casos *contrafactuales difíciles y fáciles,* están permitidos). Ambos puntos finales son arraigados en la evidencia; difieren en su tolerancia para la incorporación de *casos contrafactuales*. La mayoría de los científicos sociales prefieren explicaciones que logren un equilibrio entre *complejidad y parsimonia*. Es decir, prefieren explicaciones que están en algún lugar entre estos dos extremos.

Por ejemplo, considere el trabajo de **Mejía-Trejo (2017a)**, donde se encuentra una explicación que muestra un máximo de complejidad concluyendo:

1. Diez combinaciones causales diferentes del modelo de innovación por mercadotecnia digital (**DMIM**.*Digital Marketing Innovation Model*), compuesto por: mision-visión (**MVS**.*Mission-Vision);* propuesta de valor (**VAL**.*Value Proposition*); Mercado (**MKT**.*Market*); Metas (**GST**. *Goal Settings*); estrategia (**STG**.*Strategy*); Táctica (**TAC**.*Tactics)*; herramientas de mercadotecnia (**DMT**. *Digital Marketing Tools)*; Planeación (**PLN**.*Planning*); Desempeño (**PER**.*Performance*); Ganancias (**PRO**. *Pro itability*).

2. Ocho factores que componen al factor subyacente del modelo de competitividad nacional (México) (**NCM**. *National Competitiveness Model),* que son: agilidad (**AGY**.*Agility*); liderazgo transformador (**LDT**.*Leadership Transformer*)*;* generación de valor al consumidor (**CVG.** *Customer Value Generation;* planeación estratégica (**STP**.*Strategic Planning*)*;* guía para el cambio innovación y desarrollo contínuo (**CICD.***Guidance to Change, Innovation and Continous Development*)*;* compromise social (**SCO.***Social Commitment*)*;* bienestar e inclusion (**W&I.** *Wellness & Inclusion*)*;* Conocimiento (**KNW**.*Knowledge*).
3. Una explicación orientada a la *parsimonia,* por el contrario, se centraría en una o muy pequeña cantidad de *condiciones causales*. Por ejemplo, **Mejía-Trejo (et al., 2015)**, muestra que como innovación de la enseñanza, el modelo conceptual para aprendizaje móvil (**mL.***Mobile Learning*), se explica con tres factores: tecnología (**TECH**.*Technology*); administración de contenidos y enseñanza-aprendizaje (**C&TLM.**Contents & Teaching-Learning Management); Rol Profesor-Estudiante (**TSR.***Teacher-Student Rol*).
4. Por el contrario, una explicación que equilibra tanto a la *parsimonia* y a la *complejidad,* se centraría en distintas trayectorias a las ofrecidas por ejemplo, quizás enfocados a desarrollo político o económico y reduciendo la cantidad de salidas como resultados (o incrementarlas).

Una de las ventajas del **QCA,** es que no solo proporciona herramientas para derivar los dos puntos finales del *continuum de complejidad* / *parsimonia* sino también, proporciona herramientas para especificar soluciones intermedias. Considere, por ejemplo, la *tabla de verdad* presentada en la **Tabla 3.11**.

Tabla 3.11. Tabla de verdad con cuatro *condiciones causales* (A, B, C, D) y una salida Y

Item	Columna				
	A	**B**	**C**	**D**	**Y**
1	No	No	No	No	No
2	No	No	No	Si	¿?
3	No	No	Si	No	¿?
4	No	No	Si	Si	¿?
5	No	Si	No	No	No
6	No	Si	No	Si	No
7	No	Si	Si	No	¿?
8	No	Si	Si	Si	No
9	Si	No	No	No	¿?
10	Si	No	No	Si	¿?
11	Si	No	Si	No	¿?
12	Si	No	Si	Si	¿?
13	Si	Si	No	No	Si
14	Si	Si	No	Si	Si
15	Si	Si	Si	No	¿?
16	Si	Si	Si	Si	¿?

Nota: El resultado Y marcado con ¿?, no puede ser determinado.

Fuente: Elaboración propia.

La cual:

1. Usa **A, B, C y D** como *condiciones causales* de **Y** como resultado. Asume, como antes, que existe conocimiento del marco teórico y empírico sustantivo, que sostienen que es el presencia de estas *condiciones causales*, no su ausencia, las que están vinculadas al resultado.
2. Los resultados del análisis, salvo casos *contrafactuales* de la solución, revelan que la combinación **(A * B * ~ C)** explican a **Y**. En otras palabras, la presencia de **(A)** combinada con la presencia de **B** y la ausencia de **C** (es decir, **~ C**) explican la presencia de **Y**.

3. El análisis de esta misma evidencia, permitiendo a cualquier contrafactual que produzca un resultado *parsimonioso*, lleva a la conclusión de que **A** por sí solo representa la presencia de **Y**.
4. Concibe estos dos resultados, como los puntos finales del *continuum complejidad / parsimonia* continua, de la siguiente manera:

A * B * ~ C   A

5. Observe, que la solución que se orienta a la *complejidad* **(A * B * ~ C)** es un subconjunto de la solución que se orienta a la *parsimonia* (**A**). Esto se sigue lógicamente del hecho de que ambas soluciones deben cubrir las filas de la *tabla de verdad* con **Y** presente.
6. La solución *parsimoniosa*, también incorpora algunas de las filas con *residuales* como casos *contrafactuales* y, por lo tanto, abarca más filas. A lo largo del *continuum complejidad / parsimonia* hay otras posibles soluciones a esta misma *tabla de verdad*, por ejemplo, la combinación **(A * B).**
7. Estas soluciones intermedias se producen cuando los diferentes subconjuntos de los *residuales*, que se utilizan para producir la solución *parsimoniosa* se incorporan a la solución compleja. Estas soluciones intermedias constituyen subconjuntos de la solución más *parsimoniosa* (**A** en este ejemplo) y superconjuntos de la solución que permiten la máxima *complejidad* **(A * B * ~ C).**
8. La relación de subconjunto entre soluciones se mantiene a lo largo de la *continuum complejidad / parsimonia*. La implicación en este ejemplo, es que cualquier combinación causal que utiliza al menos algunas de las *condiciones causales* especificadas en la solución compleja **(A * B * ~ C)** es una solución válida de la *tabla de verdad* siempre que contenga las *condiciones causales* especificadas en la solución *parsimoniosa,* **A**.
9. Se deduce que hay dos intermedios válidos soluciones a la *tabla de verdad* en la **Tabla 3.11**:

(A * B)

(A * B *~ C)   (A * ~ C)   (A)

Ambas soluciones intermedias, son subconjuntos orientadas a la solución *parsimoniosa* y superconjunto de la solución compleja.

10. Los primeros **(A * B)** permiten *contrafactuales* **(A * B * C * D)** y **(A * B * C * ~ D)** como combinaciones vinculadas al resultado **Y**. El segundo enlace, es *contrafactual* **(A * ~ B * ~ C * D)** y **(A * ~ B * ~ C * ~ D)** hasta el resultado **Y**.

11. La viabilidad relativa de estas dos soluciones intermedias depende sobre la factibilidad de los *contrafactuales* que se han incorporado en ellas. Los *contrafactuales* incorporados en el primer intermedio la solución, son considerados fáciles, porque se usan para eliminar **(~ C)** de la combinación **(A * B * ~ C),** y en este ejemplo, el conocimiento existente apoya la idea de que es la presencia de **C,** no su ausencia **(~ C)**, la que está vinculada al resultado **Y**.

12. Los *contrafactuales* incorporados en la segunda solución intermedia, sin embargo, son consideradas *difíciles*, porque se utilizan para eliminar **B** de **(A * B * ~ C)** y de acuerdo con el conocimiento existente, la presencia de **(B)** debe estar relacionado con la presencia del resultado **Y**. El principio de que, solo se deben incorporar *contrafactuales fáciles*, respalda la selección de **(A * B)** como la solución intermedia óptima.

13. Observe que esta solución, es la misma que la de un investigador convencional *orientado a casos*, derivaría de la evidencia, basada en un sencillo interés en *condiciones causales* que son:

  a. Compartidas por los casos positivos (o al menos, un subconjunto de los casos positivos),

  b. Se cree que esté vinculado al resultado, y

  c. No muestra ningún caso negativo.

14. Como se ilustra en este ejemplo, incorporando diferentes *contrafactuales*, se producen diferentes soluciones. Sin embargo, estas diferentes soluciones son todas *superconjuntos* de la solución que privilegian la *complejidad* y *subconjuntos* de la solución privilegiando la *parsimonia*. Además, se ha demostrado que es posible derivar una solución intermedia óptima que solo permita *contrafactuales fáciles*.

15. Este procedimiento es relativamente simple de especificar: el investigador elimina las *condiciones causales* de la solución compleja que están en desacuerdo con el conocimiento existente, mientras se mantiene el *principio del subconjunto* que subyace en el *continuum complejidad / parsimonia*, lo que significa que la solución intermedia construida debe ser un *subconjunto* de la solución más *parsimoniosa*. Los *contrafactuales*, que son incorporados en esta solución óptima serían, relativamente rutinarios en una investigación convencional *orientada a casos*, para la misma evidencia.

Una de las grandes fortalezas de **QCA** es que todos los *contrafactuales*, tanto *fáciles como difíciles*, se hacen explícitos, como es el proceso de incorporarlos dentro de los resultados. **QCA** hace este proceso transparente y, por lo tanto, abierto a evaluación de la investigación social con orientación a la administración de la innovación.

Tomaremos el caso citado por **Mejía-Trejo (2017b)**, en el cual se sostiene que:

1. En empresas innovadoras clúster, el factor subyacente de innovación por el modelo de negocios abiertos (**Y.** Open Business Model), se da a partir de la interacción de cinco factores consideradas como variables independientes, las cuales, son:
   - **a.** Generación del Modelo de Negocio (**A.***Business Model Generation*);
   - **b.** Estrategia (**C.***Strategy*);
   - **c.** Tecnología (**S.***Technology*);
   - **d.** Nuevos Emprendimientos (**I**.*New Entrepeneurships*);
   - **e.** Orientación a la Innovación Abierta (**R**.*Open Innovation Orientation*),

2. Suponga que la *tabla de verdad* que la describe, es la **Tabla 3.12.**

Tabla 3.12. Tabla de verdad parcial de factores que intervienen en el modelo de negocios abiertos abiertos en nueve empresas innovadoras clúster

Empresa innovadora clúster	Columnas					
	1	2	3	4	5	6
	A	C	S	I	R	Y
1	Si	No	Si	Si	Si	Si
2	Si	No	No	Si	No	No
3	Si	No	No	Si	Si	No
4	No	No	No	Si	No	No
5	Si	Si	Si	Si	Si	Si
6	Si	Si	Si	Si	No	No
7	Si	Si	Si	No	No	Si
8	Si	No	No	No	No	Si

Fuente: Mejía- Trejo (2017b).

3. La **Tabla 3.12**, es la típica *tabla de verdad* pequeña de **N** casos. Muchas posibles combinaciones lógicas de *condiciones causales* (**2^5=32** filas) son representadas. Solo un puñado (**ocho filas**) tienen instancias empíricas; en consecuencia, hay una abundancia de *residuales* (**veinticuatro filas**) y, por lo tanto, gran cantidad de posibles contrafáctuales que podrían ser incorporados a la solución.
4. También se deduce que presenta *diversidad limitada* severa, dadas las muchas soluciones posibles diferentes, para esta *tabla de verdad*, dentro de los límites establecidos por los puntos finales del *continuum complejidad / parsimonia*.
5. El análisis de esta *tabla de verdad*, sin permitir la incorporación de cualquier contrafactual, produce la siguiente solución compleja:

(A * S * I * R) + (A * C * S * ~ I * ~ R) +
(A * ~ C * ~ S * ~ I * ~ R) <= (Y)

6. Estos resultados complejos, se derivan del hecho de que solo cuatro de las treinta y dos combinaciones, lógicamente posibles, muestran el resultado y ninguno de los veinticuatro restantes, se han incorporado a la solución.

7. Esencialmente, solo se ha producido una simplificación:
    a. **(A * C * S * I * R),** y
    b. **(A * ~ C * S * I * R)** se han unido para producir (**A * S * I * R)**
8. Esta solución es la más compleja posible y así establece el primer punto final del *continuum complejidad / parsimonia*.
9. Por el contrario, el uso de todos los supuestos simplificadores posibles (es decir, cualquier *contrafactual*, fácil o difícil, que ayude a producir un resultado *parsimonioso*) produce una solución dramáticamente más simple:

(~ I) + (S * R) <=(Y)

10. Esta solución, establece que la *innovación por el modelo de negocios abiertos* **Y** aún funciona, cuando no existan *nuevos emprendimientos* **(~ I),** o cuando la *tecnología y la orientación a la innovación abierta* **(S * R)** se combinan.
11. Si bien estas conclusiones, no parecen ser irracionales de esta evidencia, y mientras sean realmente concisos, van un poco en contra de las conclusiones de que un investigador convencional orientado a casos supondría.
12. Observe por ejemplo, que las cuatro combinaciones causales en la **Tabla 3.12**, relacionadas con el éxito de la innovación por el *modelo de negocios abiertos* **Y,** incluye la presencia de la *generación del modelo de negocio* **A.** Esta comunalidad del factor **A,** podría verse como un *condición causal* necesaria para el logro de **Y** y no escaparía a la atención de un ya sea de un investigador orientado a casos o de un profesional interesado en el uso de la innovación por el *modelo de negocios abiertos* **Y,** como táctica para estimular el cumplimiento.
13. Este segundo análisis, proporciona el otro punto final del *continuum complejidad / parsimonia*, que se representa de la siguiente manera:

(A * ~ C * ~ S * ~ I * ~ R)+
(A * C · S * ~ I * ~ R) +   **~ I +**
(A * S * I * R)   **(S *R)**

14. La relación de *subconjunto*, se puede observar en el hecho de que:
   a. **(A * S * I * R),** es un *subconjunto* de **(S * R),** y ambos
   b. **(A * C * S * ~ I * ~ R)** y **(A * ~ C * ~ S * ~ I * ~ R)** son subconjuntos de **(~ I)**

15. Las combinaciones causales agrupadas en cada extremo del continuum, están unidos por lógico **OR**, como se muestra en las soluciones correspondientes, denotado por **(+)**.

16. El siguiente paso es especificar soluciones intermedias y evaluarlos con respecto a los *contrafactuales* que se incorporen. Como se explicó anteriormente, una solución intermedia óptima incorpora solo *contrafactuales fáciles*.

17. Para encontrar esa solución, es necesario simplemente inspeccionar cada uno de los términos en el extremo complejo del *continuum* y determinar cuál de las *condiciones causales* separadas, si las hay, puede ser eliminado de cada combinación.

18. Considere primero la combinación **(A * S * I * R).** Las *condiciones causales* **S** y **R,** no se pueden eliminar, porque aparecen en el correspondiente término *parsimonioso*, es decir, en el otro extremo del *continuum*. Eliminar cualquiera, violaría la relación del *subconjunto*. Los únicos candidatos para ser removidos, son las condiciones **I** y **A.**

19. La generación del modelo de negocio **A** está ciertamente relacionado con el éxito de la innovación por el modelo de negocios abiertos **Y.** Por lo tanto, esta *condición causal* no debe eliminarse. Sin embargo, el hecho de que los *nuevos emprendimientos* **I** para los objetivos de la innovación por el *modelo de negocios abiertos* **Y** cambien su comportamiento, no promueve la **Y** exitosa. Por lo tanto, los *nuevos emprendimientos* **I** pueden ser eliminados de la combinación **(A * S * I * R)** debido a los inconvenientes que presenta, para cambiar el objetivo de la innovación por el *modelo de negocios abiertos* **Y** no es fundamental para el éxito de **(A * S * R),** es decir, la relación de la *generación del modelo de negocio* **A;** *la tecnología* **S** y la orientación a la *innovación abierta*, **R.**

20. Una caída de la *condición causal* de los *nuevos emprendimientos* **I**, produce el combinación intermedia **(A * S * R).** Esta combinación es un *subconjunto* de **(S * R)** y un *superconjunto* de **(A * S * I * R).**

21. A continuación, considere la combinación **(A * C * S * ~ I * ~ R).** La *condición causal* **(~ I)** (ausencia de los *nuevos emprendimientos*) no se puede descartar, porque aparece en el término *parsimonioso* correspondiente al otro extremo del continuum. Como antes, la condición *generación del modelo de negocio* **A,** no debe eliminarse porque esta condición, claramente se encuentra relacionado con el éxito de la *innovación por el modelo de negocios abiertos,* **Y.**

22. La condición *estrategia* **C** tampoco debe descartarse, porque esto también es algo que solo debería contribuir al éxito de la *innovación por el modelo de negocios abiertos* **Y.**

23. La condición *tecnología* **S** también es un factor que promueve éxito de la *innovación por el modelo de negocios abiertos,* **Y.**

24. De hecho, solo la condición **(~ R)** (ausencia de orientación a la innovación abierta) puede ser removida. Claramente, la presencia de **R** promueve éxito de la innovación por el *modelo de negocios abiertos* **Y;** es decir, las instancias de los puntos anteriores mencionadas de éxito **Y** todavía serían de éxito, de existir *orientación a la innovación abierta* **R.** Por lo tanto, esta combinación se puede simplificar como una sola condición, produciendo el término intermedio **(A * C * S * ~ I).**

25. Finalmente, considere la combinación **(A * ~ C * ~ S * ~ I * ~ R).** Nuevamente, la condición **(~ I)** debe ser retenida porque aparece en el correspondiente término *parsimonioso* y la condición **A**, también se conserva, por las razones indicadas en el análisis de las dos combinaciones anteriores.

26. La condición **(~ R)** (ausencia de *orientación a la innovación abierta*) puede eliminarse, como lo fue de la combinación previa, por el mismo motivo previsto.

27. La condición **(~ C)** (ausencia de *estrategia*) puede eliminarse, seguramente estos casos de éxito de i*nnovación por el modelo de negocios*

abiertos **Y,** aún hubieran sido exitosos si presentaran la condición de estrategia, **C.**

28. Igualmente, la condición **(~ S)** (ausencia *tecnología*) puede eliminarse de forma segura, porque solo su presencia **S** contribuye al éxito de la *innovación por el modelo de negocios abiertos,* **Y.**

29. En total, se pueden eliminar tres términos, produciendo el término intermedio (**A * ~ I**). Estos tres términos intermedios se pueden unir en una única solución:

(A * S * R) + (A * C * S * ~ I) + (A * ~ I) <= (Y)

que posteriormente se puede simplificar para:

(A * S * R) + (A * ~ I) <= (Y)

porque el término **(A * C * S * ~ I)** es un *subconjunto* del término **(A * ~ I)** y, por lo tanto, es lógicamente redundante (todos los casos de **A * C * S * ~ I** también son casos de **A + ~ I**).

30. Estos resultados, indican que dos trayectorias conducen a una *innovación por el modelo de negocios abiertos* **Y,** exitosa:

a. La *generación del modelo de negocio* **A,** combinado con la *tecnología* **S** y la *orientación a la innovación abierta* **R,** y

b. La *generación del modelo de negocio* **A,** combinado con el hecho de que la ausencia de *nuevos imprendimientos* **(~ I),** no es inconveniente.

31. La solución intermedia ahora se puede agregar al *continuum complejidad / parsimonia* continua de la siguiente manera:

(A * ~ C * ~ S * ~ I * ~ R) +		
(A * C * S * ~ I * ~ R) +	**(A * ~ I) +**	**(~ I) +**
(A * S * I * R)	**(A * S * R)**	**(S · R)**

32. Como se indicó anteriormente, la solución intermedia es un *superconjunto* de la solución más compleja y un *subconjunto* de la más *parsimoniosa*. Además es óptima porque incorpora solo *contrafactuales fáciles*, evitando el los *difíciles* que se han incorporado a los solución más *parsi-*

moniosa. La *solución intermedia* logra así un equilibrio entre *complejidad y parsimonia*, utilizando procedimientos que imitan la práctica de la investigación comparativa convencional orientada a casos. Se incluye la condición **(A)** en los resultados, según la recomendación en **Ragin (2000, pp. 105 y 254)**, para realizar las pruebas de condiciones necesarias antes de realizar, las pruebas del ser suficiente. El procedimiento *contrafactual* descrito en este capítulo, puede verse como una extensión y reformulación de las técnicas de **QCA**, una que, localiza la especificación de condiciones necesarias y suficientes dentro de un continuo de soluciones definidas por las soluciones más *complejas* y *parsimoniosas.*

33. Muchos investigadores que usan **QCA** incorporan tantos supuestos simplificadores (*contrafactuales*) como sea posible, o los evitan en total. En cambio, se sugiere lograr un equilibrio entre la *complejidad y parsimonia*, usando tanto el conocimiento empírico sustantivo como el logrado en el marco teórico de la investigación, para conducir experimentos del pensamiento, como se acaba de demostrar. **QCA** se puede usar para derivar los dos extremos del *continuum de complejidad / parsimonia*.

34. Las *soluciones intermedias*, se pueden construir en cualquier lugar a lo largo de este *continuum*, siempre que se mantenga el principio del *subconjunto* (es decir, soluciones más cercanas al extremo del *continuum de complejidad*, deben ser *subconjuntos* de soluciones más cercanos al extremo de la *parsimonia*). Una solución intermedia óptima puede ser obtenido, mediante la eliminación de *condiciones causales* individuales, que son inconsistentes tanto con el conocimiento existente de combinaciones en la *solución compleja*, que mantienen la relación de *subconjunto* con la solución más *parsimoniosa*.

Análisis contrafactual vs. investigación orientada a casos

Visto desde la perspectiva de la investigación cuantitativa convencional, la investigación comparativa *orientada a casos*:

1. Puede parecer impenetrable **(Achen, 2005)**. Los investigadores cuantitativos, saben bien que el análisis estadístico funciona mejor, cuando la cantidad de muestras de datos **N,** es grande.
2. No solo es más fácil alcanzar la significación estadística, sino que con una muestra grande de datos **N**, también pueden ahorrar a los investigadores cuantitativos, la molestia de reunir una gran cantidad de los supuestos demandantes o exigentes de las técnicas que utilizan.
3. Las violaciones de estos supuestos subyacentes, son demasiado comunes cuando el tamaño de la muestra **N** son pequeñas o incluso moderadas, como normalmente debe ser en una investigación orientada a casos.
4. Tiene la dificultad adicional, que cuando los investigadores conocen bien sus casos, tienden a construir argumentos causales combinatorios a partir de su evidencia. Desde la perspectiva de la investigación cuantitativa convencional, esta fijación en combinaciones causales, impone exigencias aún más difíciles de alcanzar para lograr evidencia en el punto de cruce.
5. También, va en contra de la lógica central de las técnicas cuantitativas más populares y utilizadas que están orientadas, principalmente, hacia la evaluación de los *efectos netos* e independientes de las variables causales, no a sus múltiples efectos combinados.
6. Sin embargo, el trabajo comparativo orientado a casos, tiene su propia lógica y rigor. Debido a que es explícitamente *configuracional,* el examen de combinaciones de *condiciones causales,* es esencial para este tipo de investigación.
7. Tal rigor, se encuentra faltante en la mayoría de las investigaciones cuantitativas, donde los casos coincidentes socavan los grados de libertad y por lo tanto, el poder estadístico. Como este capítulo muestra, sin embargo, que el estudio de combinaciones causales a menudo implica un análisis contrafactual, porque los datos sociales (por ejemplo, los orien-

tados a la administración de la innovación) que ocurren naturalmente, están profundamente afectados por la diversidad limitada, y los investigadores se comprometen proponer y realizar como casos hipotéticos, todos los *experimentos del pensamiento* (o generación de supuestos) que les sea posible.

8. Esta práctica puede nuevamente parecer sospechosa, especialmente para los investigadores cuantitativos convencionales, porque va en contra de las normas de investigación social empírica. Sin embargo, muchos de los *análisis contrafactuales*, pueden considerarse rutinarios porque implican una propuesta hipotética fácil, de los casos. Además se deben formalizar e incorporar estos *contrafactuales fáciles* en la investigación de casos cruzados dentro del marco *configuracional* del **QCA**.

9. También, se destaca una característica muy importante de la investigación social, a saber, que está construido sobre una base de fondo y conocimiento teórico, no solo de técnica metodológica. Es este conocimiento empírico sustantivo y de marco teórico de parte del investigador, que le permite evaluar la factibilidad de los casos contrafactuales.

10. En esencia, los métodos descritos en este capítulo, muestran cómo se puede entrelazar el conocimiento existente del marco teórico, con el tejido de resultados empíricos, proporcionando un método basado en el conocimiento para abordar el problema de la *diversidad limitada*.

Resumen de fsQCA para producir soluciones intermedias

Como se señaló en este capítulo, la derivación a soluciones en el continuum de complejo a *parsimonioso*, en análisis cualitativo comparativo del conjunto de *datos difusos* (**fsQCA**), es sencillo:

1. La *solución parsimoniosa*, se desprende de la designación de todas las combinaciones *residuales* (es decir, aquellas sin o con muy pocas instancias fuertes), como posibles casos *contrafactuales*.

2. La solución resultante, incorpora cualquier *combinación contrafactual* que produce una solución más simple, independientemente de si el *contrafactual es fácil o difícil*.

3. La *solución compleja*, por el contrario, no permite cualquier caso *contrafactual* y, por lo tanto, no hay supuestos simplificadores sobre combinaciones de condiciones que no existen en los datos. En efecto, la *solución compleja* define todas las combinaciones *residuales* como falsas.
4. La *solución intermedia*, no es automática, requiere la utilización del conocimiento empírico sustantiva y de marco teórico del investigador. Este conocimiento, es la base para distinguir entre *contrafactuales fáciles* (que se pueden incorporar dentro de la *solución intermedia*) y los *contrafactuales difíciles* (*que están excluidos*).
5. En **fsQCA,** la *solución intermedia* se produce junto con las *soluciones complejas y parsimoniosas* siempre que el algoritmo de *tabla de verdad* se seleccione, usando ya sea conjuntos o difusos. Una que vez la verdad la tabla está totalmente codificada por el usuario, vea **Tabla 3.13.**

Tabla 3.13. Resumen de fsQCA para producir soluciones intermedias

Item	Etapa	Descripción
1	**Revisión de la *tabla de verdad***	El investigador hace clic en el botón Análisis estándar (*Standard Analyses*) en la parte inferior de la *tabla de verdad* y: **a.** Primero, produce la solución compleja, posteriormente la solución *parsimoniosa*. **b.** Posteriormente, se consulta a los usuarios sobre sus expectativas para cada *condición causal*. **c.** Esta consulta toma la forma de un cuadro de diálogo, en el que el usuario especifica si es *la presencia o la ausencia* de la *condición causal* que debería estar vinculada a la presencia del resultado.
2	**Verificación de posibles resulatdos**	El investigador puede ingresar ya sea la presencia o la ausencia de la condición que este vinculada al resultado. Por ejemplo: **a.** Si el resultado es, evitar disminución de capacidad ceativa y una de las *condiciones causales* es, tener al menos, personal con educación de posgrado, entonces el usuario especifica que esta condición debe estar vinculada al resultado, evitando disminución de capacidad creativa, cuando se encuentre presente. Esta codificación se basaría en la conocida relación: baja educación y bajos ingresos. **b.** Por el contrario, si la *condición causal* es tener costos de operación en la introducción de innovacion de procesos, entonces el usuario especifica que esta condición debe estar vinculada a mantenerse fuera de evitar disminución de capacidad creativa, cuando ausente.

Continuación de Tabla 3.13

Item	Etapa	Descripción
2	**Verificación de posibles resulatdos**	**c.** Finalmente, si la *condición causal* es clúster geográfico, entonces el usuario haría clic en presente o ausente (*present/absent*)), ya que existe una potencial considerable de capacidad innovadora también en los clúster virtuales hoy y, por lo tanto, no hay expectativas claras de cual de estos es la mejor opción. **d.** De hecho, podría ser razonable esperar diferentes *configuraciones* de condiciones vinculadas a mantenerse evitar disminución de capacidad creativa para residentes clúster geográficos vs. clúster virtuales. **e.** Estas codificaciones de las *condiciones causales*, son las que proporcionan la base para distinguir entre *contrafactuales fáciles* y difíciles. Por ejemplo: supongamos la combinación 3 se encuentra vinculado a la prevención de evitar disminución de capacidad creativa. Suponga además, de que no hay casos de no empresa altamente innovadora, tecnología media, CEO educación de posgrado. Si el usuario no puede codificar al menos la empresa altamente innovadora como una condición que está vinculada a evitar disminución de capacidad creativa, entonces la combinación contrafactual (no empresa altamente innovadora, tecnología media, CEO educación de posgrado) se puede utilizar para simplificar la combinación existente (empresa altamente innovadora, tecnología media, CEO educación de posgrados) a tecnología media, CEO educación de posgrados. El término simplificado indicaría que la empresa altamente innovadora no es relevante para mantenerse de evitar disminución de capacidad creativa si la empresa tiene tecnología media y con CEO educación de posgrados. Sin embargo, empíricamente, esto no es cierto, para todos los casos relevantes de tecnología media y con CEO educación de posgrados al menos también son empresas altamente innovadora. En en otras palabras, sin la codificación de al menos empresas altamente innovadoras como una condición que está vinculada a evitar disminución de capacidad creativa un contrafactual difícil se incorporaría a la solución. Sin embargo, cuando está codificado correctamente, esta simplificación injustificada está prohibida. f. Asimismo, bajo estas mismas condiciones, las soluciones *parsimoniosas* pueden ser irrealmente simples, debido a la incorporación de muchas combinaciones contrafácticas (fáciles y difíciles). g. En general, se prefieren *soluciones intermedias* porque son a menudo más interpretables. Cuando la diversidad limitada es sustancial. Las *soluciones complejas* pueden serlo de forma extrema, porque ocurra una *simplificación escasa o nula.* En las soluciones intermedias, se logran un equilibrio entre *parsimonia* y *complejidad*, basadas en el conocimiento empírico sustantivo y del marco teórico propuesto por el investigador.

Fuente: Ragin (2008) con adaptación propia.

Asimetría y Equifinalidad

Ya hemos enfatizado que el funcionamiento y el modelo matemático subyacente a **QCA** se basan en supuestos fuertemente diferenciados de los que subyacen a las *técnicas estadísticas clásicas* **(Mahoney y Goertz, 2006; Ragin, 2008).** En primer lugar, en los métodos estadísticos, el incremento o disminución en una variable independiente explica el incremento o disminución en una variable dependiente *a través de una asociación directa o inversa*. **QCA**, por el contrario, descansa sobre una premisa de *asimetría*: la presencia de un cierto resultado puede estar explicado por una cierta combinación de condiciones, mientras que su ausencia puede estar explicada por condiciones explicativas completamente distintas (y no meramente por las mismas condiciones con signo contrario) **(Berg-Schlosser et al., 2009, p. 8)**. Ver **Tabla 3.14.**

Tabla 3.14. Ejemplo tabla de verdad sobre innovación por mercadotecnia digital DMIM dependiente del factor DSSME por cada caso

Empresas de innovación por mercadotecnia Casos	Columna			
	DSSMEs			DMIM (Y) Resultado
	STL	DSY	BSY	
A	1	1	1	1
D	1	0	1	1
O	1	1	0	0
P	0	1	0	0

Notas: **DMIM:** Digital Marketing Innovation Model; **DSSMEs:** Digital Skill SMEs; **STL:** Strategic Leadership; **DSY:** Digital Savvy; **BSY:** Business Savvy.

Fuente: Mejía-Trejo (2019c).

Siguiendo el ejemplo de la **Tabla 3.14**, encontraríamos que:

1. Por un lado, el resultado está presente para las configuraciones **(STL*DSY*BSY)** y **(STL*~DSY*BSY),** por lo que se puede minimizar la condición **(DSY),** que se manifiesta en estado de presencia **(DSY)** y ausencia **(~DSY).** Así, obtenemos que **(STL*BSY)** es suficiente para

explicar el resultado **(STL*BSY →Y).** Si suponemos *simetría*, deberíamos esperar que la *ausencia* del resultado **(~Y)** habría de estar causada por el signo contrario de las *condiciones explicativas* (**~STL + ~BSY →Y**). Pero ello sería erróneo. En nuestra tabla, teníamos dos casos, **O** y **P**, en los que el *resultado estaba ausente*. Para ellos, las configuraciones eran **(STL*DSY*~BSY → ~Y) y (~STL*DSY*~BSY→ ~Y).** Para *minimizar* estas configuraciones, observamos que son iguales en todo excepto en la condición **STL,** que está presente en la primera expresión y ausente en la segunda, y por tanto *puede ser eliminada*, de modo obtenemos la formulación simplificada (**DSY*~BSY → ~Y**). Como se puede observar, la explicación para la ausencia de **Y** no es la opuesta de la explicación para su presencia, sino que se trata de una explicación distinta en la que una nueva condición explicativa **DSY** interviene. No existe, por tanto, simetría.

2. En segundo lugar, y lo que es incluso más importante, **QCA** se caracteriza por la *equi inalidad:* en lugar de analizar el impacto individual de ciertas variables sobre la variable dependiente, **QCA** analiza cómo combinaciones de condiciones producen un resultado, aceptando que el mismo resultado puede estar producido por combinaciones de condiciones distintas, y por tanto que diferentes patrones causales pueden conducir al mismo resultado **(Mahoney y Goertz, 2006, p. 10)**. Esta es una ventaja relativa de **QCA** frente a otras técnicas, que descansan sobre un *supuesto de aditividad*: es decir, estas otras técnicas asumen que las variables tienen el mismo efecto de incremento sobre los resultados, con independencia de los valores que tomen otras condiciones causales relevantes **(Berg-Schlosser et al., 2009, p. 8)**.

Para comprender esto último, podemos usar un nuevo ejemplo, suponga los datos de la investigación realizada por **Mejía-Trejo (2019d)**, en las que uno de los factores (**OIN**. *Open Innovation*), es desglosado en las variables que están relacionados con las empresas Pyme epara lograr un modelo que describe la interacción de las mismas. **Ver Tabla 3.15.**

Tabla 3.15. Ejemplo tabla de verdad sobre innovación abierta (OIN) por agrupación de casos

Empresas que practican innovación abierta Casos	Columna			
	KMG	OBM	IEC	OIN (Y) Resultado
A	1	1	1	1
B	1	1	0	1
C	0	0	0	1

Notas: **OIN:** Open Innovation Efecttiveness; **KMG:** Knowledge Management; **OBM:** Open Business Model; **IEC:** Innovation Ecosystem

Fuente: Mejía-Trejo (2019c).

En este ejemplo, la *minimización* sería posible entre las configuraciones de los casos **A** y **B**, ya que son iguales en todo excepto en la condición **IEC,** que está presente en la primera y ausente en la segunda y *por tanto puede ser eliminada*. La configuración para **C,** sin embargo, no podrá ser *minimizada*, ya que es diferente a las otras dos configuraciones en más de una condición. Vea **Tabla 3.16.**

Tabla 3.16. Equifinalidad tras la minimización

Configuraciones originales	Implicaciones principales
KMG * OBM * IEC → OIN	**KMG * OBM → OIN**
KMG * OBM * ~IEC → OIN	
~KMG * ~OBM * ~IEC → OIN	**~KMG * ~OBM * ~IEC → OIN**

Fuente: Elaboración propia.

Este es un ejemplo de *equifinalidad*, donde dos configuraciones de condiciones diferentes, y por tanto *dos patrones causales distintos*, conduce al resultado **OIN**: la configuración (**KMG * OBM → OIN)** y la configuración, completamente distinta **(~KMG * ~OBM * ~IEC → OIN).**

Implicaciones principales esenciales

La obtención de los *implicaciones principales*, no es, sin embargo, siempre el último paso. En ocasiones, las expresiones pueden seguir reduciéndose. El análisis realizado por los distintos programas **QCA** proporciona normalmente las combinaciones de *condiciones más simplificadas posibles*, compuestas por lo que se denominan *implicaciones principales esenciales*. Suponga una situación en la que, tras la *minimización*, se obtienen cuatro patrones causales para el resultado **(Y)**: **(A*B*C)** , **(A*B*~C)** , **(~A*B*~C), (~A*~B*~C)**. Estos implicaciones principales, podrían reducirse aún más. Podemos imaginar, a partir de la **Tabla 3.17**, una minimización en tres pasos.

Tabla 3.17. Pasos para determinar las implicaciones principales

<table>
<tr><th>Paso 1
Configuraciones originales</th><th>Paso 2
Implicaciones principales
A</th><th>Paso 3.
Implicaciones principales
B</th></tr>
<tr><td rowspan="7">(…)</td><td>A*B*C</td><td rowspan="2">A*B</td></tr>
<tr><td></td></tr>
<tr><td>A*B*~C</td><td rowspan="2">B*~C</td></tr>
<tr><td></td></tr>
<tr><td>~A*B*~C</td><td rowspan="3">~A*~C</td></tr>
<tr><td></td></tr>
<tr><td>~A*~B*~C</td></tr>
</table>

Fuente: Elaboración propia.

De lo que se tiene:

1. En un primer paso, tenemos todas las configuraciones originales de los casos.
2. En un segundo paso, *minimizamos* siguiendo la regla de que cuando dos configuraciones son iguales en todo menos en una condición, esa condición es irrelevante y puede ser eliminada. Al hacerlo, obtenemos lo que hemos llamado *implicaciones principales* **A**, (es una etiqueta que sirve en este ejemplo, no se refiere a ningún tecnicismo). Estas primeras

implicaciones principales, sin embargo, pueden seguir siendo reducidos siguiendo la misma norma, lo que nos lleva al **paso 3**.

3. A este momento, *ya no es posible seguir reduciendo*, puesto que los *implicaciones principales* **B**, son distintos entre sí, en más de una condición. El proceso, no obstante, no está aún completo, ni son estos los que llamamos *implicaciones principales esenciales*.

La **Tabla 3.18** muestra el juego conjunto entre *las implicaciones principales* del **paso 2**, en las columnas y del **paso 3**, en las filas.

Tabla 3.18. Implicaciones principales e implicaciones principales esenciales

Configuraciones A*B*C		**Implicaciones principales paso 2**			
		A*B*C	**A*B*~C**	**~A*B*~C**	**~A*~B*~C**
Implicaciones principales paso 3	**A*B**	X	X		
	B*~C		X	X	
	~A*~C			X	X

Fuente: Elaboración propia.

Así, se tiene:

1. Como puede verse, la primera configuración del **paso 3 (A*B)** está contenida en, y por tanto cubre, dos de las configuraciones del **paso 2 (A*B*C y A*B*~C).**
2. La segunda configuración del **paso 3** (**B*~C**) cubre dos de las configuraciones del **paso 2 (A*B*~C) y (~A*B*~C).**
3. La tercera configuración del **paso 3 (~A*~C)** está contenida en otras dos configuraciones del **paso 2 (~A*B*~C)** y **(~A*~B*~C).** Lo que es importante señalar es que, entonces, para cubrir todas las configuraciones del **paso 2**, las *implicaciones principales originales*, no necesitamos todas las configuraciones obtenidas en la segunda *minimización*. La primera y la tercera configuraciones **(A*B)** y **(~A*~C)**, serían suficientes para cubrir todas las configuraciones originales, y por tanto la segunda configuración del **paso 3 (B*~C)**, *es redundante y puede ser*

eliminada. La primera y tercera configuraciones son lo que llamamos *implicaciones principales esenciales,* y serán éstos los que constituyan la solución final.

4. Finalmente, en algunos casos, la situación es incluso *más compleja*. Podemos encontrar situaciones en las que las *implicaciones principales esenciales* no vengan automáticamente dados, sino en las que se puedan hacer diferentes selecciones de *implicaciones principales esenciales*, distintas entre sí pero todas ellas capaces de cumplir con el objetivo de cubrir a las configuraciones de implicaciones originales. En este caso, el software **QCA** suele dar al usuario la posibilidad de tomar la decisión de qué configuraciones es preferible usar. Esta decisión debe venir explícitamente referida en la investigación y justificada teóricamente **(Wagemann y Schneider, 2007)**.

Dicotomización de las condiciones

El primer paso en el uso de **csQCA** es la *dicotomización* tanto de las condiciones como del resultado, sea en **(1)** cuando está *presente* y **(0)** cuando está *ausente*, y plasmarlos en una *matriz de datos dicotómica*. Vemos, en este sentido, que se propone una *simplificación de la información,* cosa que constituye un paso necesario para cualquier análisis científico. Como afirman algunos autores, cualquier investigación científica, en las ciencias sociales o más allá, en términos cuantitativos o estudios cualitativos, con un diseño experimental o no, etcétera, implica necesariamente un proceso de simplificación cuando se enfrenta a la infinita complejidad del mundo empírico (**De Meur et al., 2002, p. 120; King et al., 1994, p. 42)**. Sin embargo, recuerde, aunque sigue siendo uno de los aspectos de controversia del método, el proceso de dicotomización es uno de los más cruciales en el análisis de **QCA (Goldthorpe, 1997, p. 7; Bollen et al., 1993, p. 343)**, así que todos los datos deben dicotomizarse a partir de un umbral teórico y empírico debidamente justificado **(Rihoux y De Meur, 2009, p. 42)**. Elegir los umbrales en el proceso de *dicotomización* debe ser justificada y transparente para *reforzar la legitimidad del análisi*s **(De Meur et al., 2002,**

p. 120), ya que el análisis de los casos se fundamenta en el conocimiento en profundidad de los mismos, generalmente obtenido a través de grupos de discusión, entrevistas, revisión de archivos, bases de datos, entre otros.

Para comprender esto último, podemos usar un nuevo ejemplo, suponga los datos de la investigación realizada por **Mejía-Trejo (2019d)**, en las que uno de los factores (**OIN**. *Open Innovation*), es desglosado en las variables que están relacionados con las empresas Pyme para lograr un modelo que describe la interacción de las mismas, con los siguientes factores (**Ver Tabla 3.19**).

a. Prácticas de *administración del conococimiento* (*Knowledge Management*), **KMG.**
b. *Adopción de un modelo de negocios abiertos* (*Open Business Model*, **OBM.**
c. *Incorporación en un ecosistema de innovación* (*Innovation Ecosystem*), **IEC.**
d. *Efectividad de la innovación abierta* (*Open Innovation Efectiveness*), **OIN.**

Tabla 3.19. Matriz de datos hipotética con valores dicotómicos sobre la efectividad de la innovación abierta (OIN)

Empresas que practican innovación abierta Casos	Columna			
	Condiciones			OIN (Y) Resultado
	KMG	OBM	IEC	
A	1	1	1	0
B	0	0	1	0
C	0	1	1	0
D	0	1	0	1
E	0	0	1	0
F	1	1	0	1
G	1	0	1	1
H	1	0	0	1
I	0	0	0	1
J	1	1	1	0

Fuente: Mejía-Trejo (2019d).

Esta contiene el resultado de *efectividad de la innovación abierta* **OIN** con el valor **(1)** para la presencia del resultado y **(0)** para su ausencia. Las condiciones que intervienen **(KMG, OBM, IEC)** *también han sido dicotomizadas*. Visualizando la tabla, se pueden detectar aquellos casos con resultado **(1)** y **(0)**. Por ejemplo, de los 10 casos, el resultado **(1)** para *alta presencia de efectividad de innovación abierta* se observa en cinco casos **(D, F, G, H, I)**, y el resultado **(0)** *nula presencia de efectividad de innovación abierta* se observa en cinco casos **(A, B, C, E, J)**. Otra forma de leer la tabla es señalando que los casos **(D, F, G, H, I)**, son miembros del conjunto de empresas con *alta presencia de efectividad de innovación abierta.*

Análisis de condiciones

Se expondrá a continuación dos de las más relevantes: *necesarias y suficientes.*

Análisis de las condiciones necesarias

Un concepto importante para considerar son las *condiciones necesarias.* Una condición es *necesaria si, cada vez que el resultado está presente, la condición también está presente*. Esto quiere decir que cualquier resultado de nuestro ejemplo, *efectividad de la innovación abierta* **OIN** no puede lograrse sin la *condición* **X**; que ningún caso de presencia de **OIN** contiene la *ausencia de* **X** y que bajo la *ausencia de* **X**, la presencia de **OIN** es imposible **(Schneider y Wagemann, 2012, p. 69)**. A continuación, se discute si existe tal *condición* **X,** para la ocurrencia de **OIN**. Para **csQCA**, como para el resto de las variantes **QCA**, de nuevo, la nomenclatura tradicional para leer e identificar necesidad es la siguiente (siendo **X,** una condición cualquiera y **Y** el resultado):

$$\mathbf{X} \leftarrow \mathbf{Y}$$

Que significan: *Si* **Y***, entonces* **X** ,o, **Y** *implica* **X***,* o, **Y** *es un subconjunto de* **X.**

Como ejemplo, se va a analizar la **Tabla 3.20**

Tabla 3.20. Matriz de datos hipotética con valores dicotómicos con todos los complementas (negación) de las condiciones y una unión

Empresas que practican innovación abierta Casos	Columna							
	KMG	OBM	IEC	(~KMG)	(~OBM)	(~IEC)	(~ IEC+KMG)	OIN Resultado
A	1	1	1	0	0	0	1	0
B	0	0	1	1	1	0	0	0
C	0	1	1	1	0	0	0	0
D	0	1	0	1	0	1	1	1
E	0	0	1	1	1	0	0	0
F	1	1	0	0	0	1	1	1
G	1	0	1	0	1	0	1	1
H	1	0	0	0	1	1	1	1
I	0	0	0	1	1	1	1	1
J	1	1	1	0	0	0	1	0

Nota: Donde el signo (~) denota nula, carencia o bajo nivel del factor correspondiente.

Fuente: Elaboración propia.

Continuando con el tema de la *efectividad de la innovación abierta* **OIN**, de la siguiente forma:

1. Si la premisa de *necesidad señala que cada vez que el resultado está presente, la condición también lo está*, entonces se deben analizar aquellos casos donde el resultado **OIN,** está presente **(1)**. En nuestro ejemplo de condiciones para la *efectividad de la innovación abierta* en las empresas, se incluyen ahora los complementos (*la negación*) de las tres condiciones **(KMG, OBM, IEC),** y la unión (uso de **OR** lógico) de dos condiciones **(~ KMG + IEC).** El resultado **OIN** está presente en **(D, F, G, H, I)**.
2. Se va a comenzar con la condición **KMG**. Ésta *no pasa la prueba de necesidad* al llegar a los casos **D** e **I** puesto que el resultado está presente aunque la condición **KMG** no lo esté. Otros casos, como **(G, R, I)**, llevan a la misma conclusión para la condición **OBM**; y los casos **(D, F, H, I)** arrojan la misma conclusión para la condición

IEC. Por lo tanto, *ninguna de las tres condiciones es necesaria para la presencia de* **OIN.** Es importante destacar que esta prueba debe ser apoyada por el *análisis de necesidad* que provee el software escogido para el análisis de **QCA.**

3. El paso siguiente, es realizar la *prueba de necesidad* para los complementos (la negación de las condiciones) **(~ KMG, ~ OBM, ~ IEC).** La condición **(~ KMG** está presente en los casos **(D, l)**. Sin embargo, el resultado **OIN** está también presente en los casos **(F, G y H)**. Esto contradice la premisa de que **(~ KMG)** es necesario para la **OIN**, *por lo tanto la ausencia de esta condición no es necesaria para el resultado*. Del mismo modo, la prueba de necesidad para las condiciones **(~ OBM), (~ IEC)** revelan *que tampoco son necesarias para el resultado*.
4. Desde aquí, se puede concluir que ninguna de estas tres condiciones y sus complementos es necesaria para la *efectividad de la innovación abierta* **OIN,** en una *empresa que practica innovación abierta*, por sí mismas.
5. Una buena práctica en el *análisis de necesidad* es comenzar con el *análisis individual de las condiciones*, como ya ha sido realizado. No obstante, este análisis no finaliza aquí debido a que también se pueden crear uniones entre condiciones con el **OR** lógico. Dicha combinación de condiciones recibe el nombre de *equivalentes funcionales* (*functional equivalents*) **(Schneider y Wagemann, 2012)**.
6. La *unión* entre condiciones tiene la ventaja de crear un nuevo conjunto dentro del cual más casos tenderán a ser miembros de él y, por ende, el tamaño del conjunto se incrementará. Así, la combinación de dos condiciones podría llegar a ser un *superconjunto del resultado*.
7. Por ejemplo, con las condiciones **(~ IEC)** y **KMG** se podría crear la unión **(~ IEC + KMG)** que describe al conjunto *de empresas que practican innovación abierta* con *ausencia/ baja presencia de ecosistemas de innovación* **(~ IEC) Ó** *con presencia/ mediana/ alta de prácticas de administración de conocimiento*, **KMG.**

8. La columna **(~ IEC + KMG)** en la **Tabla 3.20** contiene los casos en los cuales esta unión está presente. Según los datos, el **OR** lógico lleva a más casos con valor **(1)** y la condición **(~ IEC + KMG)** está presente cada vez que el resultado está presente, por lo tanto, este *equivalente funcional* es una condición necesaria para la **OIN.**
9. Debido a lo anterior, es importante destacar el valor de la combinación de condiciones a través del **OR** lógico que crea la *unión* entre dos o más conjuntos, por lo que su pertenencia se determina por el *valor máximo*, entre los conjuntos de condiciones individuales. Las *uniones* de condiciones pueden ser una buena estrategia *si las condiciones individuales no logran pasar la prueba de necesidad*.
10. Sin embargo, ello debe realizarse cuidadosamente y siempre que sea justificado por el marco teórico **(Schneider y Wagemann, 2012)**. En el ejemplo de la unión de las condiciones **(~ IEC + KMG)** *no hay ninguna teoría o conocimiento empírico que respalde esta combinación* y más aún el concepto detrás de esta construcción *no se ha desarrollado*. Por lo tanto, es oportuno justificar la formación de *equivalentes funcionales* para que la interpretación del análisis cobre sentido.

Una vez realizado el *análisis de necesidad* para las condiciones que con tribuyen a la **OIN**, se procede a realizar el mismo *análisis para la negación de la* **OIN.** El paso es relativamente sencillo:

- Basta con identificar el resultado de interés con valor **(0)** y observar aquellas condiciones que sean compartidas por todos los miembros del conjunto del resultado.
- En la **Tabla 3.20**, los casos **(A, B, C, E y J)** comparten la misma condición **IEC** cuando el resultado está ausente. Así, se puede concluir que, existe evidencia que sustenta la premisa de que la presencia de *ecosistemas de innovación* **IEC** es una *condición necesaria* para la *nula/ baja presencia efectividad de la innovación abierta,* **OIN**. Esta conclusión puede ser expresada de la siguiente manera:

$$\mathbf{IEC \leftarrow (\sim OIN)}$$

Análisis de las condiciones suficientes

El análisis de suficiencia permite identificar cuándo una condición es suficiente para el resultado. *Una condición* **X** *es suficiente para el resultado, si, cada vez que esta está presente entre los casos, el resultado también está presente* **(Schneider y Wagemann, 2012, p. 75)**. La nomenclatura tradicional para expresar suficiencia es la siguiente:

$$\mathbf{X} \rightarrow \mathbf{Y}$$

Que significan: *Si* **X,** *entonces* **Y**, o, **X** *implica* **Y**, o, **X** *es un subconjunto de* **Y.**

Como ejemplo, se va a analizar la matriz de datos hipotética de la **Tabla 3.20**. Supongamos que estamos interesados en las condiciones que explican la *efectividad de la innovación abierta* **OIN** practicada en empresas innovadoras. Se estima pertenencia en tres conjuntos **(KMG, OBM, IEC)**. Como se ha explicado en la sección anterior, tanto las condiciones como el resultado están *dicotomizados* en sus valores, siendo **(1)** pertenencia *completa* y **(0)** *no-pertenencia completa*. El primer paso es identificar aquellas condiciones individuales que son suficientes para el resultado por lo que:

1. Vamos a considerar la condición **KMG** para preguntarnos si *la mediana/ alta presencia de la práctica de la administración del conocimiento en una empresa innovadora*, es suficiente para la ocurrencia de *efectividad de innovación abierta* **OIN**. Si es así, entonces cada vez que esta condición está presente en una empresa la efectividad de la innovación abierta, también debiera estar presente.
2. Del mismo modo, ningún caso con la condición **KMG** puede estar relacionado con la *ausencia* del resultado **(~ OIN).** A efectos prácticos, primero, se van a observar solo aquellos casos donde la condición **KMG** tiene valor **(1)**. Como puede apreciarse, **KMG** tiene valor **(1)** para los casos **(A, F, G, H y J),** y la condición **KMG** está relacionada también a la *ausencia* del resultado en los casos **(A, J)**.

3. Esta información es concluyente para afirmar que la *mediana / alta presencia práctica de la administración del conocimiento* **KMG** *no es una condición suficiente para la alta/mediana presencia de la efectividad de la innovación abierta,* **OIN.**
4. La condición **OBM,** está presente en los casos **(A, C, D, F, J).** Si la *mediana / alta presencia de la adopción de negocios abiertos* de una empresa innovadora *fuera suficiente para la ocurrencia de la efectividad de la innovación abierta* **OIN**, entonces el valor de la condición tendría que ser **(1)** para los cinco casos. Sin embargo, los casos **(A, C, J)** no son miembros del conjunto de empresas innovadoras, por lo tanto, **OBM** *no es una condición suficiente para el resultado.*
5. La condición **IEC** está presente en los casos **(A, B, C, E, G, J)** y, como se puede observar, esta condición está relacionada con la *ausencia* del resultado en los casos **(A, B, C, E, J).**
6. Por lo tanto, se puede concluir que **IEC** *no es una condición suficiente para el resultado.*
7. Como para el análisis de necesidad, la prueba de suficiencia también debe aplicarse a los *complementos de las condiciones:* **(~ KMG), (~ OBM), (~ IEC).** Si las condiciones **(~ KMG)** y **(~ OBM),** son individualmente suficientes para la **OIN**, entonces *el resultado debe estar presente en todos los casos que incluyan cada una de dichas condiciones*.
8. Sin embargo, esto no es el caso, debido a que *ambas condiciones no cumplen con el criterio de suficiencia para la* **OIN**:
    - **a.** El caso **B,** para la primera condición, así como **(C y D)** para la segunda, proporcionan evidencia de la *inexistencia de suficiencia* para ambas condiciones.
    - **b.** Por su parte, **(~ IEC)**, está presente en los casos **(D, F, H, I,)** y en todos estos casos el resultado está también presente **OIN**.
9. Por lo tanto, se puede concluir que la *ausencia de incorporaciones a ecosistemas de innovación* **IEC,** de una empresa innovadora, es una condición suficiente para la **OIN.**

10. Estas conclusiones, si bien es cierto que son relevantes para una interpretación, *no responden en su totalidad a la pregunta sobre qué condiciones son suficientes para la* **OIN.**
11. Por ejemplo, al analizar particularmente el caso **G,** se observa que este no contiene la ausencia de **IEC** y que, sin embargo, hay presencia de **OIN**, pese a la ausencia de la condición que ha sido detectada como suficiente para el resultado. Esto puede entenderse como la *constatación de la existencia de distintas trayectorias que llevan al resultado,* y como evidencia de que la **OIN** aún puede ocurrir con la *presencia de la incorporación a ecosistemas de innovación* **IEC** *de una empresa innovadora*.
12. En consecuencia, el paso siguiente a realiza, es la identificación de aquellas *condiciones suficientes* que cuentan para aquellas instancias de **(OIN)** que no son explicadas por la *nula/escasa/baja/ausencia de la incorporación a ecosistemas de innovación* **(~IEC)** de una empresa innovadora.

En el *análisis de suficiencia, un paso más allá del análisis de condiciones individuales es la prueba de suficiencia de conjunciones*, por ejemplo **(KMG * OBM)**. La ventaja de este proceso es detectar casos en los cuales la combinación conjunta de dos condiciones puede explicar el resultado. En el caso de **(KMG * OBM)** se puede leer como *las empresas innovadoras que son caracterizados por la combinación de una mediana / alta presencia de prácticas de administración de conocimiento y la mediana / alta presencia de negocios abiertos tienen mediana / alta presencia de efectividad en innovación abierta*.

Para efectos del análisis, se ha creado una matriz de datos **Tabla 3.21**, con las conjunciones bajo el **AND** lógico de condiciones individuales, que conjuntamente, serán *suficientes* para el resultado.

Tabla 3.21. Matriz de datos hipotética con valores dicotómicos con conjunciones

Empresas que practican innovación abierta Casos	Columna							
	KMG	OBM	IEC	(KMG*OBM)	(KMG*IEC)	(OBM*IEC)	(~ OBM*KMG)	OIN Resultado
A	1	1	1	1	1	1	0	0
B	0	0	1	0	0	0	0	0
C	0	1	1	0	0	1	0	0
D	0	1	0	0	0	0	0	1
E	0	0	1	0	0	0	0	0
F	1	1	0	1	0	0	0	1
G	1	0	1	0	1	0	1	1
H	1	0	0	0	0	0	1	1
I	0	0	0	0	0	0	0	1
J	1	1	1	1	1	1	0	0

Nota: Donde el signo (~) denota nula, carencia o bajo nivel del factor correspondiente.

Fuente: Elaboración propia.

Como ejemplo se han incluido tres conjunciones **(KMG * OBM), (KMG * IEC) y (OBM * IEC).** Como se observa en la tabla:

1. Las conjunciones no tienen muchos miembros con valor **(1)**, lo cual puede ser explicado por la regla del valor mínimo del **AND** Lógico, que crea la *intersección* entre dos o más conjuntos, por lo que su pertenencia se determina por el *valor mínimo*, entre los conjuntos.
2. Según los datos, *ninguna de las tres conjunciones creadas cumplen el criterio de suficiencia*, sin embargo, la creación de conjunciones no termina aquí, puesto que aún se pueden crear conjunciones entre las tres combinaciones y sus complementos.
3. A modo de ilustración se ha insertado la columna **(~ OBM * KMG)** en la tabla para su análisis. Como puede observarse, dicha conjunción está presente en los casos **(G, H)**, donde el resultado **OIN,** también está presente.

4. Por lo tanto podemos concluir que la conjunción de una *nula/ escasa/ baja presencia de adopción de modelos de negocios abiertos* **(~ OBM) Y** *la mediana/alta presencia de prácticas de administración de conocimiento* **KMG,** es una condición que cumple el criterio de *suficiencia* para que el resultado de *efectividad en la innovación abierta* **OIN**, ocurra.
5. No hay que olvidar que este ejemplo es hipotético y solo tiene sentido para efectos ilustrativos de este análisis. La evidencia empírica de estudios de inmigración es por supuesto investigaciones consolidadas tanto teórica como empíricamente y no es precisamente el objetivo de este estudio ponerlos en discusión.
6. Volviendo a nuestro análisis, se había identificado como *condición suficiente* para la inmigración a **(~ IEC)** una condición que cubre los casos de **(D, F, H, l).** Ahora el *equivalente funcional* **(~ OBM * KMG)** cubre los casos **(G, H).**
7. Por lo tanto en esta sección se puede concluir que *hay al menos dos caminos que son suficientes* para que ocurra la *efectividad de la innovación abierta en una empresa innovadora*:

   (~ IEC) + (~ OBM * KMG) → OIN

   Esta solución puede leerse como: *la nula/ escasa/ baja/ ausencia de incorporación a ecosistemas de innovación* (casos **D, F, H I**) *o la combinación de un nulo/ escaso/ baja presencia adopción de modelos de negocios abiertos* ***Y*** *una mediana/ alta presencia de prácticas de administración de conocimiento (casos* ***G , H****) es suficiente para que la efectividad de la innovación abierta se presente en una empresa innovadora.*

En resumen, una causa se define como *necesaria*, si debe estar presente para que ocurra un resultado. Una causa se define como *suficiente*, si por sí misma puede producir un resultado determinado. Esta distinción es significativa solo en el contexto de perspectivas teóricas. Ninguna causa es *necesaria*, por ejemplo, independiente de una teoría que la especifique como

una causa relevante. Ni ser *necesaria* ni la *suficiente*, existen independientemente de las teorías que proponen las causas. El *ser necesario y ser suficiente*, generalmente se consideran juntas porque todas las combinaciones de las dos, son significativas. Así:

1. *Una causa es necesaria y suficiente*, si es la única causa que produce un resultado y es singular (es decir, no es una combinación de causas).
2. *Una causa es suficiente pero no necesaria,* si es capaz de producir el resultado, pero no es la única causa con esta capacidad.
3. *Una causa es necesaria pero no suficiente*, si es capaz de producir un resultado, en combinación con otras causas y aparece en todas esas combinaciones.
4. *Finalmente, una causa no es necesaria ni suficiente*, si aparece solo en un *subconjunto* de las combinaciones de condiciones que producen un resultado.

En total, hay cuatro categorías de causas, formadas a partir de la tabulación cruzada, de la presencia / ausencia de ser suficiente vs. la presencia / ausencia de ser necesaria.

La aplicación típica de **QCA** (*nítida o difusa*) da como resultado, una declaración lógica que describe combinaciones de condiciones que, son suficientes para el resultado. La lista de combinaciones, puede ser o no exhaustiva, es decir, puede no explicar todas las instancias del resultado. Es una buena idea examinar, tanto el ser necesario como el ser *suficiente* de las condiciones individuales, antes de iniciar el análisis de ser *suficiente* en las combinaciones de condiciones. Esto se puede hacer, al observar diagramas de dispersión del resultado, por cada condición y tomar nota de cuáles son *cuasi superconjuntos* (es decir, ser necesaria) y cuáles son *cuasi subconjuntos* (es decir, ser *suficiente*) (ver también análisis de *subconjuntos/ superconjuntos*).

Minimización

El proceso de *minimización Booleana,* es aquel que *permite reducir la complejidad del conjunto de datos inicial de la tabla de verdad*. La *minimización* funciona del siguiente modo: cuando dos filas en la *tabla de verdad* muestran exactamente la misma configuración (incluso para el resultado), excepto en una condición explicativa que en una fila está presente y en otra ausente, esta puede ser considerada *irrelevante o redundante* en la explicación de los resultados.

El carácter restrictivo de la lógica combinatoria, parece indicar que su enfoque es complejo; este no es el caso. Existen reglas simples y directas para simplificar la *complejidad*, para reducir expresiones y formular declaraciones Booleanas más concisas. La más fundamental de estas reglas es:

> *"Si dos expresiones Booleanas difieren en una sola condición causal, pero producen el mismo resultado, entonces la condición causal que distingue las dos expresiones, puede considerarse irrelevante y puede eliminarse para crear una expresión combinada más simple".*

Esencialmente, sobre la mencionada regla de minimización y retomando el caso de la **Tabla 3.2**.

Le permite al investigador tomar dos expresiones Booleanas que difieren en un solo término y producir una expresión combinada. Por ejemplo, el **(E * ~ C * ~ S)** y el **(E * C * ~ S),** que producen en los resultados **M,** difieren solo en la **(C).** Todos los demás elementos son idénticos. La regla de minimización establecida anteriormente permite el reemplazo de estos dos términos con una sola expresión más simple: **(E * ~ S).** En otras palabras, la comparación de estas dos filas: **(E * ~ C * ~ S)** y **(E * C * ~ S),** como un todo, indican que en casos de **(E * ~ S),** el valor de la **(C)** es *irrelevante*. La condición de **C** puede estar presente o ausente; **M** seguirá ocurriendo.

Tabla 3.2. Tabla de verdad hipotética de innovación

Empresa	Columnas				
	Efectividad de la ejecución de la innovación tecnológica	Cambio del director de la ejecución de la innovación tecnológica	Insatisfacción del cliente del producto innovado	Pérdida de posición de mercado	Número de casos (frecuencia)
	E	C	S	M	F
1	0	0	0	0	9
2	1	0	0	1	2
3	0	1	0	1	3
4	0	0	1	1	1
5	1	1	0	1	2
6	1	0	1	1	1
7	0	1	1	1	1
8	1	1	1	1	3

Fuente: Elaboración propia.

La lógica de esta simple reducción de datos, es paralela a la lógica del diseño experimental. Solo una *condición causal*, la **C,** varía y no se detecta ninguna diferencia en el resultado (porque tanto el **E * ~ C * ~ S** como el **E * C * ~ S** son casos que causan **M**). Según la lógica del diseño experimental, la **(C)** es *irrelevante* para **M** en presencia de **(E * ~ S)** (es decir, *se deben mantener estas dos condiciones constantes*). Por lo tanto, el proceso de minimización Booleana imita la lógica del diseño experimental. Es una operacionalización directa de la lógica de la comparación social-científica ideal.

Este proceso de minimización lógica, se lleva a cabo de forma ascendente, hasta que no sea posible una mayor reducción gradual de las expresiones Booleanas. Considere nuevamente, los datos sobre la pérdida del *posicionamiento de mercado* **M**, presentados anteriormente. Cada una de las filas, con una causa presente y dos ausentes, se pueden combinar con filas con dos causas presentes y una ausente, porque todas estas filas tienen el mismo resultado **M** y cada par, difiere en una sola *condición causal* (ver **Tabla 3.22**):

Tabla 3.22. Ejemplo de minimización.
Caso: una causa presente con dos causas ausentes

Item	Columnas		
	Expresión 1	Combina con expresión 2	Para producir
1	E * ~ C * ~ S	E * C * ~ S	E * ~ S
2	E * ~ C * ~ S	E * ~ C * S	E * ~ C
3	~ E * C * ~ S	E * C * ~ S	C * ~ S
4	~ E * C * ~ S	~ E * C * S	~ E * C
5	~ E * ~ C * S	E * ~ C * S	~ C * S
6	~ E * ~ C * S	~ E * C * S	~ E * S

Fuente: Elaboración propia.

Del mismo modo, cada una de las filas con dos causas presentes y una ausente, se pueden combinar con la fila con las tres presentes (ver **Tabla 3.23**)

Tabla 3.23. Ejemplo de minimización.
Caso: dos causas presentes con una causa ausente

Item	Columnas		
	Expresión 1	Combina con expresión 2	Para producir
1	E * C * ~ S	E * C * S	E * C
2	E * ~ C * S	E * C * S	E * S
3	~ E * C * S	E * C * S	C * S

Fuente: Elaboración propia.

Es posible una mayor reducción. Tenga en cuenta, que los términos reducidos en la primera ronda, se pueden combinar con los términos reducidos de la segunda, para producir expresiones aún más simples (ver **Tabla 3.24**).

Tabla 3.24. Ejemplo de minimización.
Caso: una causa presente con una causa ausente

Item	Columna		
	Expresión 1	Combina con expresión 2	Para producir
1	E * ~C	E * C	E
2	E * ~ S	E * S	E
3	~ E * C	E * C	C
4	C * ~ S	C * S	C
5	~ E * S	E * S	S
6	~ C * S	C * S	S

Fuente: Elaboración propia.

Este proceso, de minimización produce al final :

(E) + (C) + (S) = M

Es cierto que esto era obvio, por la simple inspección de la *tabla de verdad*, pero el problema presentado, fue elegido por su simplicidad. El ejemplo ilustra directamente, las características clave de la minimización Booleana. Es de abajo hacia arriba. Busca identificar conjuntos de condiciones cada vez más amplias (es decir, combinaciones más simples de *condiciones causales*) para las cuales un resultado es, verdadero. Y es como un experimento, en su enfoque en pares de *configuraciones*, que difieren solamente, en una sola causa.

Análisis de ocurrencias del caso

A continuación se realizará un explicación sobre la ocurrencia y no *ocurrencia del caso de innovación abierta* **OIN** hasta el momento estudiada, así como su evaluación.

Análisis ocurrencia de caso: innovación abierta (OIN)

Retomando nuevamente de los datos incluidos en la **Tabla 3.1.**

Tabla 3.1. Que muestra una *tabla de verdad* hipotética, sobre el modelo de innovación abierta de Mejía-Trejo (2019d)

Items	Columnas					
	Condiciones			Resultado		Casos
	KMG	OBM	IEC	OIN	(~OIN)	
1	1	1	1	0	1	A, J
2	0	0	1	0	1	B, E
3	0	1	1	0	1	C
4	0	1	0	1	0	D
5	1	1	0	1	0	F
6	1	0	1	1	0	G
7	1	0	0	1	0	H
8	0	0	0	1	0	1

Nota: Donde el signo (~) denota nula, carencia o bajo nivel del factor correspondiente.

Fuente: Mejía-Trejo (2019d).

Se tiene:

1. El primer paso es crear una expresión Booleana de todas aquellas filas que estén vinculadas con el resultado a explicar, en este caso, **OIN**.
2. Las filas con resultado **(1)** son **(4, 5, 6, 7, 8)** y, a modo de ejemplo, la **fila 4** puede ser escrita como **(~KMG * OBM * ~IEC)**, la cual es una conjunción que en una *tabla de verdad* recibe el nombre de *expresiones primarias o primitivas*.
3. En expresión Booleana, las filas con resultado **(1)** pueden ser expresadas del siguiente modo:

Fila 4 + Fila 5 + Fila 6 + Fila 7 + Fila 8

(~KMG*OBM~IEC) + (KMG*OBM*~IEC) + (KMG*~OBM*IEC) + (KMG*~OBM*~IEC)+(~KMG*~OBM*~IEC) → OIN

4. Cada una de estas expresiones primarias, es considerada como *condición suficiente para el resultado de interés*. La expresión Booleana que se ha creado es la *expresión más compleja para expresar suficiencia* en una *tabla de verdad*.

5. Por lo tanto, el paso siguiente es aplicar el *principio de minimización* mencionado anteriormente: las condiciones redundantes pueden ser omitidas y las filas restantes pueden ser unidas en una conjunción de condiciones más simplificada.
6. Al aplicar este principio a la **fila 7 (KMG*~OBM*~IEC)** y **fila 8 (KMG*~OBM*~IEC)** se puede ver que son idénticas excepto por el valor de la condición **KMG**. Esto quiere decir que se está ante la presencia de dos configuraciones vinculadas al resultado de interés y que difieren en una sola condición.
7. Si la regla señala que dicha condición lógicamente irrelevante debe ser quitada, entonces **KMG** es eliminada de dichas filas, lo que conduce a reescribir la formulación del siguiente modo:

Fila 4 + Fila 5 + Fila 6 + Filas 7 y 8

(~KMG*OBM~IEC) + (KMG*OBM*~IEC) + (KMG*~OBM*IEC) + (~OBM*~IEC) → OIN

8. La información de las **filas 7 y 8** indica que el *nulo/ escaso/ ausencia/ baja adopción de los modelos de negocios abiertos* (**~OBM**) *en combinación* **con** *el nulo/ escaso/ ausencia/ baja incorporación a los ecosistemas de innovación* **(~IEC),** constituye una *condición suficiente* para la mediana/ alta presencia de la *efectividad de innovación abierta* **OIN**.
9. Seguidamente, analizando las **filas restantes 4 y 5**, se observa que ellas difieren solo con respecto al valor de la condición **KMG**, por lo tanto este puede ser removido del análisis.
10. Las *filas 4 y 5* pueden interpretarse como que la *mediana/alta presencia de adopción de modelos de negocios abiertos.* **OBM,** *combinado con el nulo/ escaso/ ausencia/ baja incorporación a los ecosistemas de innovación* **(~IEC)**, es una *condición suficiente* para *la mediana/ alta presencia de la efectividad de innovación abierta* **OIN** y que puede expresarse de Ia siguiente manera:

Filas 4 y 5 + Fila 6 + Filas 7 y 8

(OBM~IEC) + (KMG*~OBM*IEC) + (~OBM*~IEC) → OIN

11. Siguiendo con el proceso de minimización de las expresiones primarias o primitivas restantes se observa que la condición **OBM** difiere entre la **fila (4 y 5) y (7 y 8)**. Al quitarla del análisis, la expresión simplificada queda como:

Filas 4 y 5 + Fila 6 + Filas 7 y 8

(~IEC) + (KMG*~OBM*IEC) + (IEC) → OIN

12. Esto indica que *el nulo/ escaso/ ausencia/ baja incorporación a los ecosistemas de innovación* **(~IEC)** de una empresa innovadora es una condición suficiente para *la mediana/ alta presencia de la efectividad de innovación abierta* **OIN**.
13. La fórmula más simplificada sería la siguiente:

Filas 4, 5, 7 y 8 + Fila 6

(~IEC) + (KMG*~OBM*IEC) → OIN

14. Para seguir minimizando y obtener la *solución parsimoniosa*, es importante analizar cuál es el papel de la condición **IEC,** cuando está combinado con **(KMG*~OBM).**
15. La condición **IEC**, *no se requiere para la parsimonia* debido a que la conjunción (**~OBM*IEC)** incluye ambas expresiones primitivas **(KMG*~OBM*IEC)** (**fila 6**) y **(KMG*~OBM*~IEC)** (**fila 7**).
16. Por lo tanto, al decir que **(KMG*~OBM)** es una condición suficiente para la **OIN**, también se está afirmando que tanto **(KMG*~OBM*IEC)** y **(KMG*~OBM*~IEC)** *son su icientes para el resultado de interés*.
17. Más aún, como estas dos expresiones difieren solo en el estatus de **IEC**, entonces esta condición puede ser removida del análisis.
18. Así la solución que se obtiene es:

(~IEC+~OBM*KMG) → (OIN)

Las diferentes soluciones obtenidas en el proceso de minimización comparten diversas características como la *equivalencia*, el hecho de expresar la información contenida en la tabla de verdad, carecer de contradicciones y reflejar un estado resumido de la información recolectada.

Según las buenas prácticas del método de **QCA**, la solución final a escoger en el *proceso de minimización debe ser justificada cuidadosamente sobre base empírica, teórica y lógica.*

Análisis no ocurrencia de caso: innovación abierta (OIN)

Para el análisis de **(~OIN)**, el proceso es el mismo al ya descrito anteriormente:

1. Se deben considerar las filas de la **Tabla 3.1** que se vinculan con el resultado **(0)**, es decir, **(1, 2, 3).** Del mismo modo, se debe tomar en cuenta el *análisis de necesidad* realizado a **(~OIN)**, en el apartado de *análisis de condiciones necesarias* de este capítulo, el cual señalaba la condición **IEC,** como *necesaria* para el resultado de interés.
2. Para realizar el *análisis de suficiencia* de **(~OIN)**, en lugar de repetir manualmente la operación seguida en el subapartado anterior, se aplicará directamente el *algoritmo* **Quine-McCluskey** que es utilizado en el software **fsQCA**.
3. Este lleva a la siguiente solución:

(IEC*~KMG) + (IEC*OBM) → (~OIN)

4. Esta fórmula puede ser reescrita al factorizar la condición **IEC** como:

IEC(OBM*~KMG) → (~OIN)

5. La narrativa del resultado obtenido puede leerse, a modo de ejemplo, como que la *baja/nula/escasa/ausencia de efectividad de innovación abierta* **(~OIN)** ocurre en empresas innovadoras que tienen *mediana/alta presencia de incorporación en ecosistemas de innovación* **IEC, Y** que presentan una *baja/nula/escasa/ausencia de prácticas de administración de conocimiento* **(~KMG) O**, alternativamente, que tienen *mediana/alta incorporación en ecosistemas de innovación* **IEC** que presentan *mediana/alta adopción de modelos de negocios abiertos* **OBM.**

6. Como puede apreciarse, en las dos vías de la solución la condición **IEC** está presente, confirmando su carácter *necesario* para la ausencia de **(~OIN)**. Sin embargo, esta conclusión, aunque muchas veces pueda parecer evidente, no es tan sencilla de confirmar cuando una *tabla de verdad* contiene *residuales y filas contradictorias* (aunque no es el caso aquí, dado que no hay *residuales*).
7. En esta sección, se debe tener en cuenta la identificación y tratamiento de los residuales y de las *filas contradictorias* para evitar errores en cuanto a la *presencia y ausencia de condiciones*. En la segunda parte de este capítulo se presentará una guía sobre cómo tratar este asunto con un ejemplo empírico.
8. Finalmente, la mínima expresión obtenida confirma la *ausencia de simetría* entre las soluciones para ambos resultados y deja en evidencia que las causas de la alta y la baja **OIN,** son fenómenos cualitativamente diferentes y que requieren, por tanto, de *análisis de componentes distintos*.

Evaluación de resultados

Para interpretar los resultados después del *proceso de minimización*, se puede seleccionar la *solución de interés* (*compleja, parsimoniosa o intermedia*) o alguna de sus partes que cumplan con el *umbral de consistencia*. Estas pueden ser analizadas a través de dos procesos:

1. El retomo a los casos, y
2. La evaluación de la teoría **(Ragin, 1987)**.

Ambos procesos, cuando se unen, deben *establecer mecanismos causales*, es decir, aquellas *fuerzas causales* que existen entre la configuración y el fenómeno bajo estudio (la conocida *caja negra*). Por ejemplo:

1. En la solución *parsimoniosa* de nuestro ejemplo sobre **OIN,** se observa que **(~IEC + KMG * ~OBM)** contribuye a la **OIN**.

2. Se podría estudiar la primera ruta **(~IEC)**, lo que conduce a seleccionar los casos cubiertos por este **(D, F, H, I)**, y desenmarañar los mecanismos que contribuyen a la **OIN** en los mismos.
3. Luego, se procedería a evaluar la teoría vinculada a esta ruta de acuerdo con la literatura seleccionada en el estudio, es decir, evaluar cuán relevante es la ausencia de gobiernos conservadores en un país receptor dentro de la teoría de la **OIN**.

   (~IEC) → Mecanismo causal → OIN

4. Cuando nos enfrentamos a *equifinalidad*, el investigador puede ir y venir entre las distintas rutas que contribuyen al resultado y analizar sus correspondientes casos para establecer la existencia de rasgos distintivos que pudieran explicar mucha equifinalidad **(Wagemann y Schneider, 2007; Ragin y Rihoux, 2004)**.
5. En nuestro ejemplo existen dos rutas que explican la inmigración en una empresa innovadora y dos que explican la **OIN** en el caso **H**. Encontrar la *singularidad* de cada camino para explicar **H** es una adecuada forma de interpretación de los resultados.
6. Del mismo modo, el investigador puede poner énfasis más en la teoría que en los resultados y viceversa, dependiendo de cuáles sean las preguntas y objetivos de investigación establecidos.
7. Por ejemplo, si se interpreta la segunda ruta que contribuye a la **OIN**, el investigador se podría preguntar si la historia causal es la misma para los casos **(G, H)** planteando, por ejemplo, las siguientes preguntas:
   - **a.** ¿Cuál es el papel de la *ausencia de incorporarse a los ecosistemas de innovación* **IEC** en la historia del caso **H**?
   - **b.** ¿Qué distingue a los casos de la primera ruta y qué distingue a los casos de la segunda?
   - **c.** ¿Cuál es la narrativa detrás del hecho que de acuerdo con la segunda ruta, la *ausencia de adopción de modelos de negocios abiertos* **OBM** combinada con la presencia de altas prácticas de administración de conocimiento **KMG** han contribuido a la *alta efectividad de la innovación abierta* **OIN** en los casos **(G y H)**?

d. ¿En qué medida los casos de **(G, H)** o, de otra parte, **(D, F, H, I)**, son comparables?
e. ¿Por qué **(~IEC),** como condición clave, es suficiente para explicar la ocurrencia de **OIN**?

Preguntas similares, pueden ser formuladas para explicar la negación del resultado sobre la base de las mismas estrategias ya presentadas. Estas interrogantes orientan la interpretación hacia un *análisis causal* acerca de los mecanismos que producen o no el resultado de interés. De hecho, las conclusiones de un estudio con **QCA** tienen diferentes metas, en comparación con un procedimiento *estadístico clásico* de rechazar o confirmar una teoría. El método de **QCA** evalúa teorías pero también identifica el contexto en el que operan los casos. Una cierta teoría, representada por una combinación de condiciones, podría estar asociada con el resultado esperado en un contexto específico, pero no en otro. Como señala **Ragin (1987, p. 123)**.

> *"...el producto final típico de un análisis Booleano, es una declaración de los límites de las variables causales identificadas con diferentes teorías, no su rechazo o aceptación mecánica..."*.

Así, este proceso de ir y volver a los casos no es en vano, puesto que proporciona un aumento en el conocimiento de los mismos y un flujo en la formulación de nuevas hipótesis orientadas a la teoría.

Cómo investigar con QCA

Para realizar investigación con **QCA**, se sugieren tres puntos básicos que se deben considerar y tener presentes, con el fin de confirmar si es la herramienta adecuada a aplicar en su diseño de investigación, que ha saber, son:

1. **QCA** es compatible con varios diseños de investigación, en particular los *casos más semejantes con resultados diferentes* (**MSDO**. *Most*

Similar cases with Difierent Outcomes) en los llamados estudios con **N** *mediana o intermedia*, a explicar más adelante.

2. **QCA** es un diseño de investigación comparativo, por lo que la actividad de la comparación, está en el centro de sus principales operaciones. No es solo seleccionar casos diferentes, sino de pensar la investigación en términos comparados. Esto afecta los conceptos empleados y, en consecuencia, en el redactado como pregunta de investigación. Por ejemplo, preguntarse por qué la crisis financiera global ha impactado con menos dureza a alguno de los países reportados en la *Global Innovation Index* **(GII, 2019)** demanda de la actividad comparativa entre países.
3. **QCA** es sólido, dado que requiere de una clara secuencia de decisiones, y cada una de ellas debe ser discutida y presentada de forma transparente.

Debemos recordar que un *diseño de investigación*, es la organización general de un corpus de conocimiento teórico y empírico que se someten a la articulación de las decisiones y operaciones prácticas en su transcurso y que permiten alcanzar los objetivos de la investigación de la mejor manera posible. Es general, los objetivos se encuentran muy vinculados a las preguntas de investigación que con **QCA**, están motivadas por la explicación de los resultados (*outcome-driven*). De hecho, **QCA** se estructura en torno al resultado, excepto si se emplea para la construcción de tipologías.

Restrictivamente, existen autores que sostienen que la *investigación comparada* se reduce a dos tareas principales: la selección de los casos para dar respuesta a las preguntas de investigación y la selección de las variables, lo que indica construir algún tipo de *modelo causal*. Por lo anterior, se sugiere una secuencia de cuatro pasos:

1. Definición de los casos.
2. Decidir cuántos casos serán incluidos en el análisis empírico.
3. Decidir qué casos serán seleccionados.
4. Lograr el diálogo con los casos durante el análisis **QCA**.

El carácter secuencial de estos pasos, puede ser alterado en el proceso de investigación con **QCA,** como de ocurre frecuentemente con cualquier investigación con casos múltiples, por lo que pueden producirse varias iteraciones y saltos entre pasos (Rihoux y Lobe, 2009). La investigación con **QCA** no debe realizarse de manera lineal y mecánica. Por ejemplo, problemas que aparecen en el desarrollo analítico (*paso 4*) pueden desembocar en una reconsideración de la definición de los casos (*paso 1*).

Determinando los casos

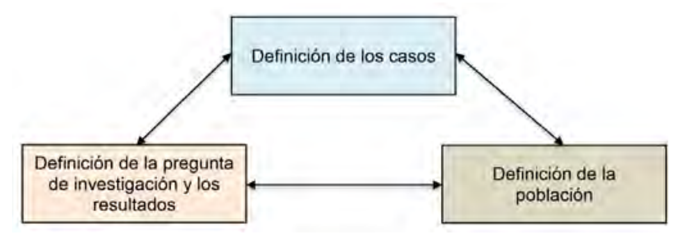
El Dr. Ragin encontró un camino intermedio entre los enfoques cualitativos y cuantitativos **(Ragin 1987)**, **QCA** es más bien, un enfoque orientado a los casos (*case-oriented approach*) **(Ragin, 2009b; Rihoux, 2006; Rihoux y Lobe, 2009)**. Es por tal motivo, que el concepto de *configuración* se vuelve central en las operaciones de **QCA**, dado que cada *configuración es una narrativa codificada de un caso* con un cierto número de *condiciones* que conllevan a algún *resultado de interés*. Dado que **QCA** requiere de alguna forma de *codificación numérica* para obtener la *tabla de verdad* a fin de procesar estos datos numéricos, con base a un sistema de software, las *tablas de datos* en **QCA** no deben leerse como tal (*data points*), como se haría en análisis estadísticos con **N** grande. Cada caso, debe ser considerado como un sistema en sí mismo **(Przeworski y Teune, 1970)**. Con esto, el investigador inicialmente debe tener en cuenta además, tres aspectos principales, rara vez evidentes **(Medina et al., 2017)** a definir:

a. Por caso, definir el marco teórico conceptual y empírico sustantivo que los sostiene,

b. Cuál es la pregunta de investigación, con énfasis particular, al resultado a explicar; y

c. Cómo se circunscribe y relacionarla con la población potencial de los casos.

La interconexión entre estos tres aspectos es considerablemente alta, como se plasma en la **Figura 3.1.**



Fuente: Medina (et al., 2017).

Esto quiere decir, que la secuencia en la que se toman decisiones sobre estos tres temas no está perfectamente establecida; todo depende del contexto de la investigación. En algunas situaciones, la formulación de la pregunta de investigación, la cual contiene el resultado a explicar, así como una primera definición del caso (o los casos), pareciera ser relativamente obvia. Por ejemplo, en un análisis inicial que partan de la siguiente observación, de acuerdo con el *Global Innovation Index* **(GII, 2019)** en algunos países de América del Sur tuvieron una prevalencia del factor: subíndice de entrada de innovación (*Innovation Input Subindex*) relativamente baja, de 2015 a 2019. Por contraposición, se produjo un incremento muy considerable, durante el mismo periodo, en los países de América del Norte. A partir de esto, es posible formular una sencilla pregunta de investigación: ¿qué combinación de factores podrían explicar los diferentes *subíndices de entrada de innovación* para los países de América del Sur?

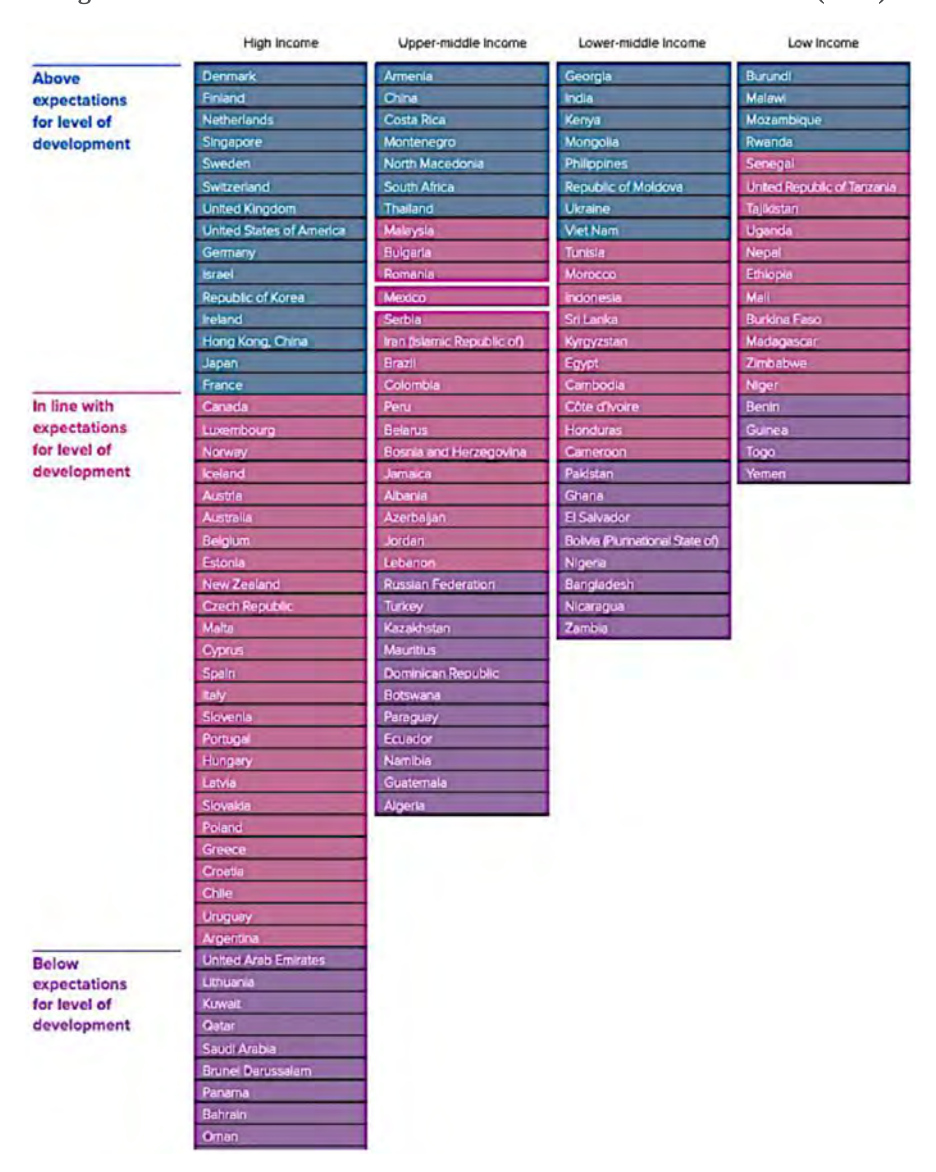
Observe en este ejemplo, que los tres temas mencionados se encuentran cubiertos :

1. El *resultado* se encuentra claramente definido (la evolución de los *subíndices de entrada de innovación*);
2. La naturaleza y el nivel de los casos también está resuelta (países); y
3. La población de los casos logra estar perfectamente circunscrita (países de América del Sur).
4. Se observa además que la pregunta de investigación: expectativas sobre combinaciones de factores concuerda con el uso de **QCA** para los análisis posteriores.

En este primer ejemplo, se cubre la totalidad de la población de los casos (países de América del Sur), lo que no es frecuente en situaciones reales. En otras situaciones, los casos y la población están dados de antemano por la propia naturaleza de la investigación. Por ejemplo, un especialista que pretenda evaluar el desarrollo de una política de innovación de la región occidente de México tendrá que cubrir la totalidad de los estados que la componen (toda la población) si se trata de un encargo de las autoridades federales. Los casos estarán definidos en un nivel agregado (por sector industrial, por estado, etcétera). El *resultado* también está de algún modo definido con antelación: por ejemplo, para operacionalizar la *evaluación del índice de innovación nacional* se recurre a indicadores socioeconómicos oficiales. En situaciones extremas, el punto de partida es incluso menos claro. Por ejemplo, el mismo *Global Innovation Index* **(GII, 2019)**, prevé que a pesar de que los factores (llamados subíndices en este caso) se diseñan para hacer clasificaciones rigurosas, existan *casos de frontera* (*boundary-cases*) es decir, países que no pueden ser plenamente comparados con otros a pesar de estar por ejemplo en la clasificación como México de altos-medianos ingresos (*Upper-middle income*). Para lograr su clasificación, se interpone otro criterio: (*Above/ In line/ Below expectations for level of development*). Ver **Figura 3.2.**

Esto fuerza a los diseñadores de los índices a una redefinición de la población de los casos ya que se descubren variaciones en las variables e indicadores de los mismos subíndices. Para un investigador que tome de referencias estos reportes (y otros), deberá tomar en cuenta y ser consciente de la interacción entre la *definición de resultado* (presente en su pregunta de investigación) y la *definición del caso* (presente en la caracterización que haga del mismo, de acuerdo con su contexto de espacio y tiempo).

Figura 3.2. Caso de México de acuerdo al Global Innovation Index (2019)



Fuente: GII (2019).

Este último ejemplo muestra que, a menos que el investigador trabaje con un conjunto de casos predefinidos a la par que se encuentren extremadamente bien circunscritos, la tarea de definición de los casos es laboriosa y no del todo evidente. De esta forma, al observar un fenómeno social que involucre a la innovación, la pregunta que tiene que hacerse es: ¿de qué caso, formalmente definido como concepto de nuestro marco teórico, forma parte este caso, observado empíricamente? En investigaciones estadísticas convencionales con una muestra **N** grande, los casos son genéricos y dados, por ejemplo, cada uno de los **400** encuestados en una encuesta de formas de innovación. En investigaciones con un número pequeño o intermedio de muestras de casos, de lo contrario, así como para un gran número de casos no catalogados como individuos tales como: organizaciones, administraciones, estados, empresas, procesos políticos, etcétera, las fronteras empíricas entre los casos se vuelven menos sólidas y cuestionables. De esta forma, el investigador necesita literalmente, construir empírica y conceptualmente sus casos.

Así, los casos tienen que encontrar su propia definición dentro de un área de homogeneidad o campo de investigación **(Berg-Schlosser y De Meur, 2009)** que fije las fronteras dentro de las cuales los casos serán, primero, definidos y, segundo, seleccionados. En nuestro caso de ejemplo, de la evaluación del índice de innovación nacional, el área de homogeneidad es el país (México) donde las acciones y políticas de innovación han tendido lugar. Todos las empresas (casos potenciales) operan dentro del mismo entorno institucional: México, caracterizado por un sistema de innovación distribuido en 32 estados. Con ello, el investigador puede estar seguro de la definición de la población (de empresas dedicadas practicar innovación) a partir de la cual poder obtener una muestra mínima de 400 casos (empresas) con las que realizar una encuesta. La seguridad radica en el hecho de poder controlar por un gran número de factores potenciales de varianza externa **(Mejía-Trejo (2019a, 2019b)**, como son los factores contextuales que podrían afectar la práctica de la innovación. Frente a situaciones con **N** pequeñas menos sencillas, digamos que la comparación de empresas dedicadas a servicios altamente innovadores basados en internet

(casos potenciales), con un resultado a explicar que sea el éxito financiero más problemático en este tipo de empresas, surgen a primera vista los siguientes retos:

1. **1.** Primero, ¿cómo se define una empresa de servicios innovadora en internet, cuáles son sus propiedades esenciales, cómo deben tratarse los casos de frontera, por ejemplo una empresa de servicios que se autodenomina innovadora pero que difiere en su práctica en algunos aspectos?
2. **2.** Segundo, ¿son todos las empresas de servicios innovadores basadas en internet comparables sin atender su propio tamaño, el tamaño de la región, estado o país que atienden? ¿No sería más lógico comparar una empresa de servicios innovadores basados en internet de la región occidente vs. los que operan al norte de México que con los de la Ciudad de México.
3. **3.** Tercero, ¿es equiparable la realidad empírica de las empresas de servicios innovadores basados en internet, financieramente hablando con los del norte, occidente y centro de México?

Como consecuencia de estas y otras más interrogantes, para estar seguros, el área de *homogeneidad* debería contener probablemente solo a los estados del occidente, norte y centro de México, y a su vez, tengan un tamaño determinado. Dadas estas y otras más dificultades, una estrategia útil en investigaciones con una **N** pequeña o intermedia es empezar con un caso típico de un fenómeno concreto. Esto permite definir las propiedades de lo que sería el tipo ideal de lo que es un caso. Por ejemplo, *Uber* como taxi en cualquier momento y lugar para transportación de personas, *Uber eats* como caso completo e ideal de una empresa que se dedica a la distribución de comida; o *Rappi* también de distribución de comida, e incluso de servicios adicionales vía internet. Desde el primer momento en que las propiedades definitorias centrales del caso típico quedan perfectamente determinadas, a la vez de que se formula un concepto que corresponde con esas propiedades, entonces se ha definido un caso, es decir, la *unidad*

de observación. Este paso facilita enormemente la observación de otros ejemplos empíricos, de otros casos que son lo bastante similares como para permitir la comparación.

Como se observa, los casos son, por lo tanto, *construcciones*. Los casos en sí, no existen como tales: los casos se circunscriben y son definidos por el investigador. Una política de innovación en México existe, pero ¿es el país lo que consideramos caso, o una institución pública, un sector tecnológico, un servicio basado en internet, un CEO, o un proceso estatal, o una política pública, o una región o clúster dentro de ese estado o país? ¿Cómo se encuentra este caso definido en términos temporales? En definitiva, cuanto más clara sea la definición empírica y conceptual del caso, con mayor claridad se podrá delimitar en tiempo y espacio, así como más segura será la comparación entre casos y el análisis QCA en particular **(Medina et al., 2017)** La definición del caso requiere siempre de una serie de decisiones arbitrarias y, aun así, siempre se encuentran *casos fronterizos y difíciles* cuya pertenencia a la población de casos pareciera dudosa. Más adelante, en el proceso de desarrollo de **QCA**, el investigador puede verse forzado a añadir o suprimir unos cuantos casos fronterizos.

El alcance de la cantidad de casos

Una vez definidos los casos, el siguiente paso es decidir el número de casos que se quieren o se necesitan incluir en el análisis empírico. Se deben considera tres consideraciones para tomar la mejor decisión al respecto:

1. *El tamaño de la población.* En este punto, las investigaciones que estudian organizaciones o grandes entidades (Estados, regiones, clústeres, sectores de la industria, etcétera) como casos se encuentran muchas veces ante la posibilidad de analizar empíricamente la población entera. Es posible analizar los **32** Estados de México, o los clúster de innovación de la región norte del país, o las **400** empresas que realizan innovación por negocios electrónicos de Guadalajara y Monterrey, o el sector de empresas dedicadas a desarrollo de software en la Ciudad de México.

Lo mismo sucede cuando se analizan individuos como casos: los **30** ganadores de algún concurso de gobierno que asigna recursos para desarrollo de la innovación tecnológica, o los **10** empresas que son reconocidas en algún contrato para desarrollar polos de innovación como en el estado de Jalisco. En la medida en que sea posible obtener suficiente información sobre los casos, no hay inconveniente en lidiar con toda la población de casos en un análisis QCA, sobre todo porque esta no requiere de inferencia estadística. De hecho, cubrir la población entera de casos es una estrategia de investigación muy poderosa, pues hacer generalizaciones a partir de una muestra (o selección) de casos puede llegar a ser arriesgado **(Medina et al., 2017)**.

2. *Los aspectos prácticos.* El investigador debe resolver y convencerse de una respuesta convincente a la siguiente pregunta: ¿hasta qué punto es posible añadir más casos (*anchura*) y todavía mantener el suficiente conocimiento sobre ellos (*profundidad*)? **QCA** se fundamenta en el conocimiento sobre los casos (*case-based knowledge*) o, como **Ragin (1987)** lo plantea, **QCA** requiere de suficiente intimidad con cada caso. Añadir casos, por lo tanto, depende mucho del acceso a información secundaria (monografías sobre el caso, informes, publicaciones, información encontrada en internet, etcétera), de la calidad de dicha información y de los recursos con los que cuente el investigador para conseguir material adicional para cada nuevo caso. **QCA** exige información sólida para cada caso puesto que el investigador debe proporcionar información para cada una de las celdas de la *tabla de verdad*. Si la información es incompleta o defectuosa, entonces está totalmente justificado, antes de emplear **QCA**, la exclusión del caso del análisis empírico.
3. El equilibrio entre *anchura y profundidad* está estrechamente relacionado con las decisiones que se toman con respecto a los objetivos de la investigación. Así, se tienen las siguientes reflexiones:
   a. Si se lleva a cabo un análisis exploratorio, dando prioridad al conocimiento dentro de los casos (*within-case understanding*) (con preguntas que busquen el *cómo*) y a la elaboración de teoría, es recomendable mantener un número pequeño de casos.

b. De lo contrario, si el foco se pone en un *análisis causal o explicativo*, es decir, *preguntar por qué*, y en probar teorías (*theory-testing*), a la vez que el investigador cuenta con bastantes condiciones que contribuyan potencialmente a la explicación de ocurrencia (o no ocurrencia) del resultado, entonces es preferible ampliar el número de casos. La razón es bien sencilla: un mayor número de casos genera mayor diversidad empírica **(Ragin, 2006)** lo que refuerza la seguridad de las afirmaciones causales después de un **QCA**. Con ello es también mucho más fácil generalizar a otros casos colindantes (*neighbouring cases*).

c. La aplicación de **QCA** en situaciones de una **N** grande es perfectamente posible. Hay una gran cantidad de este tipo de aplicaciones. En algunas de estas aplicaciones, **QCA** se utiliza para reanalizar datos de encuesta o censales **(Blackman y Dunstan, 2010)** donde los casos son individuos. **QCA** es diferente al orientarse menos a los casos. No hay duda de que en este tipo de diseños no es posible llegar a conocer en profundidad cada uno de los casos; sin embargo, el investigador debe familiarizarse lo suficiente con los diferentes tipos o categorías de casos. Por ejemplo, si se está realizando datos de una encuesta sobre el uso de la nube (*cloud*) en el desarrollo de innovaciones de productos en México, el investigador debe tener una imagen (o preconcepción) clara de cómo deberían operar las diversas categorías típicas del desarrollo de la innovación de productos (digamos, diseño de productos con conexión a internet), así como sería ideal que el investigador fuera capaz de conseguir acceso a información mucho más cualitativa (por ejemplo, entrevistas) que pudiera proporcionar un mayor conocimiento de los propios casos.

Cómo seleccionar los casos

Las investigaciones con **QCA**, de casos múltiples, requieren una clara definición del resultado (y del caso individual) que debe estar presentada en los principios de la investigación, porque esta información será necesaria para saber:

- Qué casos se convierten en candidatos
- Hasta qué punto se buscan (o no) casos con diferentes resultados.

Una primera necesidad, será aclarar si se opta por una muestra o por una selección de casos, con las principales notas:

a. El muestreo es considerada una operación mecánica que permite al investigador de forma más o menos automática, siguiendo unos cuantos parámetros estadísticos, seleccionar una muestra dentro de una población bien definida. Se podría decir que el muestreo es algo así como una siembra (con ayuda de una máquina).

b. La selección de casos no es una operación tan mecánica como el muestreo, sino consecuencia de decisiones tomadas por el propio investigador. La selección de los casos, es la recogida manual de lo expuesto en el punto de muestreo.

Los principales contrastes se encuentran en la **Tabla 3.25.**

De hecho, se sugiere que: cuando el investigador trabaje con una muestra N mediana o intermedia, con la pretensión de analizar los datos con QCA, la mejor estrategia será optar por la selección de casos, no la del muestreo. Esto está en línea con una de las principales características de QCA como enfoque: el investigador debe tener un papel activo, con aportaciones de peso, a lo largo del proceso de investigación.

Tabla 3.25. Diferencias entre la estrategia de muestreo y la selección de casos

	Selección de casos	Muestreo
Definición de los casos	Las fronteras y la definición de los casos están sujetas a debate	Los casos vienen dados
Delimitación de la población	Las fronteras y la definición de la población están sujetas a debate	La población viene dada
Objetivo	Seleccionar casos (o las categorías de los casos) más apropiados. No hay un criterio *a priori* en términos de máximo o mínimo del mínimo de casos	Seleccionar casos que son suficientemente numerosos y que cubren suficientemente las diferentes categorías (no casos individuales) dentro de la población, con tal de obtener una muestra representativa que permita la inferencia estadística de la muestra de la población
Operación	Depende del conocimiento y la dedicación del investigador	Técnica (operación estadística)
Criterio de calidad	La calidad es siempre discutible y sujeta a las propias decisiones del investigador (*anchura* vs. *profundidad*). Tener más casos no es sinónimo de obtener mejores resultados	Cuanto más grande sea la muestra, mejor será. La inferencia estadística será mucho más precisa entonces.
Anchura de la población	Puede tener sentido seleccionar todos los casos de una población para maximizar la diversidad	No tiene sentido mantener todos los casos dado que la inferencia estadística permite el ahorro

Fuente: Medina (et al., 2017).

Anterior al **QCA**, la estrategia de selección de casos más habitual, al menos cuando se emplea para evaluar hipótesis (derivadas de la teoría), proposiciones o conjeturas, es la conocida, como *casos más semejantes con resultados diferentes* (**MSDO**. *Most Similar cases with Difierent Outcomes*), del fenómeno a analizar. Esta estrategia está basada en el razonamiento lógico y experimental, originario en la obra de **Mill (Berg-Schlosser y De Meur, 2009),** y en la lógica del *diseño de sistemas más similares* (**MSSD**. *Most Similar Systems Design*) **(Przeworski y Teune, 1970)**.

El principio del diseño de *casos más semejantes con resultados diferentes* (**MSDO**. *Most Similar cases with Different Outcomes*) es bastante simple, si:

a. El investigador ha definido un área de hegemonía controlando por varianza exógena, y
b. Dentro de esta área de hegemonía el investigador selecciona un cierto número de casos con variación en el resultado; y
c. El investigador obtiene información sobre factores (*condiciones*) potenciales, que a su vez son variables entre los casos que ejercen algún impacto sobre el resultado; y
d. Existen argumentos teóricos y/o conocimiento sustantivo empírico del propio investigador para pensar que estas condiciones tienen algún impacto sobre el resultado, entonces,
e. La comparación sistemática, o sea, *múltiples comparaciones por pares* de los casos, permitirán identificar las condiciones que realmente logran tener algún impacto sobre el resultado.

El primer paso, es garantizar que se están seleccionando los casos que realmente ofrecen variación en el resultado. Un ejemplo, tomando un resultado *dicotómico* con valores altamente diferentes es, si el investigador quiere explicar la supervivencia de algunas empresas antes de la implementación de políticas de innovación y la actualidad dentro de un periodo determinado, entonces se tendrá que examinar también un cierto número de casos, cuyo resultado fuera diferente a la supervivencia: *la no supervivencia* o el colapso de las empresas **(Berg-Schlosser y De Meur, 2009)**. De hecho, esta variación en el resultado es necesaria para evaluar cualquier teoría o hipótesis causal. Desde una perspectiva **QCA**, *no existe el supuesto de simetría causal*, por ejemplo, al encontrar una explicación para la presencia del resultado, no se puede automáticamente deducir la explicación para su ausencia ya que podría tratarse de una explicación completamente diferente. En este ejemplo hipotético empírico, la situación es sencilla debido a que se tienen en consideración (como la

población total) las empresas de cierto sector de la industria donde se supone existan una clara variación en el resultado de los casos.

Ahora, suponga una situación hipotética empírica más compleja, cuando los contrastes en el resultado no son tan evidentes, y el investigador no está muy interesado en examinar la negación del resultado **(Y = 0),** no es necesario encontrar alguna variación en el resultado. Una regla general es que se necesita al menos un cuarto de los casos tomando algún valor (0 y 1) del resultado. Lo mismo sucede para las condiciones, si es que no hay alternativa, por lo que se debe considerar:

1. Es perfectamente legítimo, seleccionar casos por el valor de su resultado, así, seleccionar por el resultado es en principio una mala práctica en investigaciones con **N** grande, pero cumple todos los requisitos cuando se trabaja con **N** pequeñas centradas en los casos. Es básicamente resultado de seguir la lógica del pensamiento contrafactual, que se realiza en el momento de comparar dos casos con resultados diferentes **(Tarrow, 2010)**: uno empieza buscando un caso típico de resultado positivo **(1)**, y después busca un caso contrario **(0)** u opuesto aunque todavía comparable, como un caso típico de empresa inalterado en la misma región.

2. El segundo paso, es asegurarse de que existe variación entre todas las condiciones que forman parte del modelo. El objetivo del investigador debe ser, asegurarse un máximo de heterogeneidad sobre un número mínimo de casos. En otras palabras, asegurarse una *rica diversidad entre los casos*.

3. Aquí es donde la teoría adquiere relevancia. Normalmente, las *condiciones causales* se originan de la *teoría*, mientras que el investigador se encarga de *formular hipótesis* direccionales que vinculen cada condición con el resultado. La forma habitual sería: se espera que un alto valor para la condición **A** lleve a un alto valor para el resultado **Y**. Si la investigación consta de **25** casos, **15** con un alto valor en el resultado **Y** y **10** con un valor bajo en el resultado **(~ Y)** y la condición **A,** muestra un valor bajo para los **25** casos, entonces, por definición, esta condición es una *constante*. Si es así, por lógica, la condición **A** no puede explicar

la variación en el resultado, se vuelve una variable de control, y es oportuno *sacarla del modelo.*

Claro está, considerar que una *condición* es esencialmente una *constante* depende del juicio del investigador tras preguntarse: *¿existe variación significativa o no?* Por ejemplo, en el momento de explicar el colapso o la supervivencia de las empresas antes y después de una política de innovación en cierto periodo, una condición con pleno significado hubiera podido ser: la *presencia* (o *ausencia*) de una fuerte confianza por participar en dicha política. Algunos investigadores argumentarían que todas las empresas participantes mostrarían, de acuerdo con el contexto, un fuerte sentimiento de progreso, lo que convertiría dicha condición en una *constante*. Sin embargo, podría darse el caso de que algunas empresas fueran caracterizadas por un altísimo nivel de confianza ,lo que podría servir para reflejar variación. En cualquier caso, de nuevo, se requiere la implicación, o la justificación mediante el marco teórico que soporte el investigador.

Finalmente, la *selección de las condiciones* se enfrenta a una dificultad frecuente en las ciencias sociales particularmente en el fenómeno de la innovación: la abundancia de condiciones causales potencialmente significativas, por ser un campo marcado por múltiples, cuando no competidoras, teorías. Esto se vuelve un problema para investigaciones con **N** intermedia, en tanto que el número de condiciones debe mantenerse bajo, pues, de lo contrario, surge un problema de diversidad limitada: muchas variables, pocos casos. Contamos con algunas estrategias para superar esta dificultad **(Berg-Schlosser y De Meur, 2009)**:

- Por un lado, la estrategia más segura es evaluar por separado una versión corta de la teoría, por ejemplo, incluyendo solamente las *condiciones más relevantes*.
- Por otro lado, una estrategia más elaborada es empezar por un modelo abierto alimentado por múltiples teorías, y después reducirlo gradualmente hasta obtener unas pocas condiciones que se adapten al diseño **MSDO**. Para hacer esto se puede recurrir al procedimiento **MSDO/**

MDSO, que es costoso en términos de tiempo, pero muy riguroso y transparente.

QCA: buenas prácticas vs. dificultades

Todo lo anterior sugiere que, antes de dar comienzo al análisis, el investigador habrá:

a. Definido claramente los casos;
b. Tomado decisiones precisas sobre la profundidad de los casos, la anchura entre ellos y la generalización;
c. De asegurarse de que existe suficiente homogeneidad del universo de la investigación;
d. Definido claramente los contornos de la población de casos, para asegurar comparabilidad entre ellos; y
e) haber seleccionado un número de casos que permitan la máxima diversidad entre el resultado y las condiciones.

El diálogo entre el investigador y los casos debe continuar en la parte analítica, en esa parte que emplea el software para la sintetización para obtener la *tabla de verdad* y la *minimización* en donde se obtienen las *soluciones más parsimoniosas*. **QCA** no está diseñado como un mero ejercicio de utilización mecánica de las herramientas de un software; el uso del software siempre se desarrolla en diálogo con el conocimiento teórico y sobre los casos que tenga el investigador.

Una visión general a un proyecto de investigación con **QCA** se caracteriza por tres fases:

a. *Exploración* **QCA**, principalmente para adquirir conocimiento sobre los casos, definir la población, los casos y el modelo, así como conseguir datos relevantes para crear la matriz de datos;
b. *Análisis* **QCA** (incluido el uso de software), empezando por la *dicotomización o calibración*, seguido de varios pasos técnicos y, finalmente,

la obtención de soluciones **QCA** *más o menos parsimoniosas*, lo que Ragin llama el *momento analítico*; y

c. *Interpretación* **QCA** incluyendo la generalización, así como la *vuelta a los casos* y a la teoría (Rihoux y Lobe, 2009). La puesta en marcha de estas tres fases *puede no ser lineal* **(Mahoney y Rueschemeyer, 2003)**; estas se entrelazan, hay saltos e iteraciones entre ellas, tal y como argumentaron **Rihoux y Lobe (2009)** con la idea del embudo de la complejidad (*funnel of complexity*).

La recomendación práctica importante a estas alturas es que, más allá de los pasos técnicos que requieren el uso del software, *el investigador debe volver a los casos y al marco teórico* a lo largo de diferentes pasos del análisis **QCA**. En particular, la *vuelta a los casos* es particularmente útil e instructiva en las siguientes etapas:

a. *La dicotomización o calibración,* dependiendo del uso de **csQCA**, **mvQCA** o **fsQCA**, del resultado y las condiciones. Prestar atención a los casos para localizar con precisión los *puntos de corte* es un buen ejercicio. Sin embargo, en las ciencias sociales, los puntos de corte puramente teóricos, en el fenómeno de la innovación, siempre son objeto de disputa. Lo mismo sucede con las condiciones próximas a una *dicotomía natural*. Por ejemplo, si los casos son **30** individuos CEO y el género fuera una de las condiciones, uno de los individuos puede ser transexual (tras haber pasado por una operación de cambio de sexo, de mujer a hombre). El investigador debe decidir cuál es el tratamiento que recibe dicho individuo bien como hombre (postoperación) bien como mujer (preoperación), o simplemente se decide sacar este caso del estudio. Algo similar sucede cuando se trata de *percepciones*.

b. *Solucionar configuraciones contradictorias*, que es probablemente uno de los retos más importantes en **QCA**. Este tipo de configuraciones albergan casos con resultados iguales a **(1)** y a **(0)**, a la vez que muestran los *mismos valores para las condiciones*. Una de las mejores estrategias para resolver *configuraciones contradictorias* es la de reexaminar

con detalle, de manera *cualitativa*, los casos que forman parte de dichas configuraciones para, después, reconsiderar cómo se han operacionalizado algunas condiciones y el resultado. Esta estrategia permite, además, descubrir que uno o varios de los casos incluidos en una configuración contradictoria es de hecho un *caso de frontera* que encaja perfectamente con la población anteriormente perfilada.

c. *El uso de los residuales durante la minimización*. Mientras que la principal estrategia para decidir qué remanentes son consistentes es la argumentación teórica; ¿qué *residuales* caben dentro de las expectativas teóricas?, también tiene sentido realizar *experimentos mentales* (*thought experiments*) *sobre casos no observados*, a partir del conocimiento del propio investigador sobre el campo de estudio y los casos observados.

d. *Solucionar los supuestos simplificadores contradictorios* (*Contradictory Simplifying Assumption*s) **(Yamasaki y Rihoux, 2009)** por medio del conocimiento previo sobre los casos para asignar un *valor en el resultado* a los *residuales* lógicos que son problemáticos **(Vander-brought y Yamasaki, 2004)**, y con ello, obtener soluciones **QCA** mucho más significativas.

e. *Decidir entre los términos de las soluciones* **QCA**. Frecuentemente, el software **QCA** produce rutas diferentes y alternativas, aunque parecida, por ejemplo, compuestas por el mismo número de condiciones, que podría explicar el resultado para un caso o para un conjunto de casos. En esta situación, un conocimiento profundo de los casos puede servir para decidir a favor de una combinación de condiciones algo más factible y significativa.

La idea general es que cada paso técnico del análisis **QCA** (usando el software) puede ser mejorado si se establece un diálogo con el conocimiento que el investigador tenga de los casos. Esto puede consumir mucho tiempo, pero el beneficio se traduce en soluciones **QCA** altamente robustas y fundamentadas. Naturalmente, al afrontar los últimos pasos de **QCA**, en particular la interpretación individual de cada caso y en su conjunto, el investigador debería retomar a los casos para traducir la solución **QCA**

en una narrativa (véase capítulo 1). Esta narrativa puede ofrecerse más o menos en términos de ruta causal (*causal path*), condiciones antecedentes/consecuentes, proceso, condiciones más o menos distantes o cercanas (catalizadoras), etcétera. Todo esto requiere conocimiento de los casos. Probablemente la mejor forma de plantear estas interpretaciones sería consultar a especialistas, o incluso los directamente afectados (decisores públicos, líderes de las organizaciones analizadas, etcétera), como se ha hecho en algunas aplicaciones **QCA** en investigaciones de gestión pública.

Como se ha discutido en este capítulo, **QCA** es compatible con varios diseños de investigación, desde **N** pequeña a **N** grande, aunque la forma más frecuente de utilizar **QCA** es el diseño de *casos más semejantes con resultados diferentes* (**MSDO**. *Most Similar cases with Difierent Outcomes*) en investigaciones con **N** mediana o intermedia. No obstante, al margen del nivel de variación del resultado y del número de casos empíricos considerados, algunas buenas prácticas son transversales y permanentes. Probablemente las cuatro mejores prácticas son:

a. Primero, tomar los casos y el conocimiento sobre ellos de forma seria.
b. Segundo, no usar **QCA** como algo mecánico, pulsando cualquier botón existente en el software. Es importante reflexionar y reconsiderar todas las decisiones tomadas con antelación de una forma iterativa.
c. Tercero, ser sistemático en el modo en que se llevan a cabo todos los pasos prácticos antes, durante y después del **QCA**.
d. Cuarto, ser transparente en la presentación y discusión de todas las decisiones tomadas.

En términos de elecciones prácticas, a pesar de que *casos más semejantes con resultados diferentes* (**MSDO**. *Most Similar cases with Different Outcomes*), es el diseño de investigación más frecuente en **QCA**, no existe un único diseño para todas las investigaciones. El investigador puede tomar todo tipo de decisiones sobre los objetivos de su investigación, los impedimentos y oportunidades, quizás optando por el pragmatismo, y de hecho sobre sus propias preferencias entre investigaciones con

N pequeña centradas en el conocimiento profundo de los casos o investigaciones con **N** grande enfocadas al análisis entre los casos y generalizaciones más amplias.

QCA y su combinación con otros métodos de análisis

QCA al ser de naturaleza dual como *técnica holística y analítica*, tiene la cualidad de combinarse con otros métodos, ya sean cualitativos o *cuantitativos*. Esto es considerado una *buena práctica* en lo que a innovación se refiere **(Rihoux et al., 2009, pp. 170-172; Schneider y Wagemann, 2012)**.

En un estudio rápido, realizado en **Web of Science**, con el uso de las palabras clave como **Títulos**: **QCA** or **CSQCA** or **MVQCA** or **FSQCA**

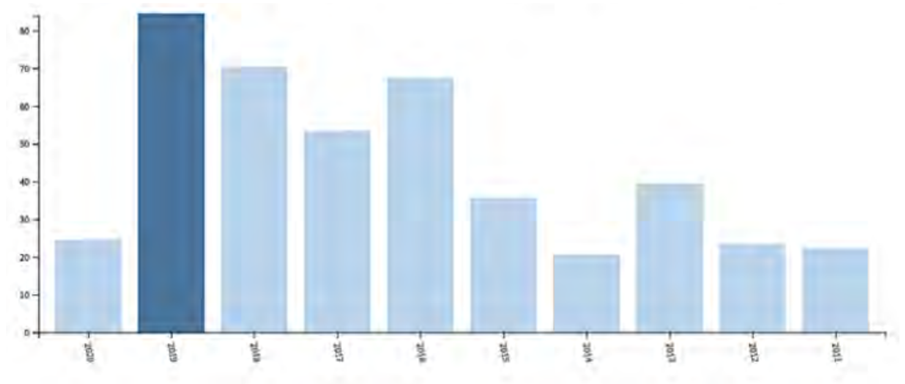
En un periodo de **1980-2020**, se obtuvieron **607** documentos, como se muestra en las **Gráficas 3.1 y 3.2.**

Gráfica. 3.1. Publicaciones que relacionan al título QCA, csQCA, mvQCA y fsQCA por especialidad. Mapa de árbol: 607 documentos

Fuente: Web of Science Clarivate (2020).

Se resaltan las publicaciones: **158** de ingeniería eléctrica y electrónica (engineering electrical), **104** de sistemas cardiacos y cardiovasculares (*cardiac cardiovascular systems*), **78** de nanociencias y nanotecnología (*nanoscience nanotechnology*) y **64** de negocios (*business*).

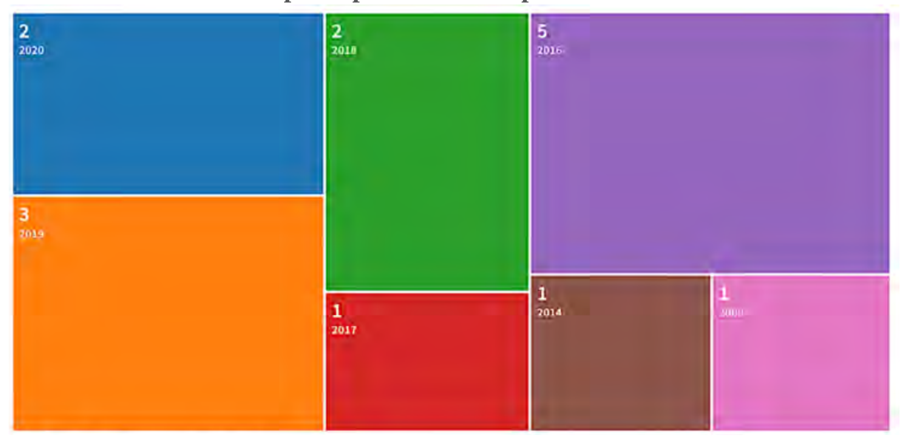
Gráfica. 3.2. Publicaciones que relacionan al título QCA, csQCA, mvQCA y fsQCA por especialidad. Diagrama de barra: 607 documentos



Fuente: Web of Science Clarivate (2020).

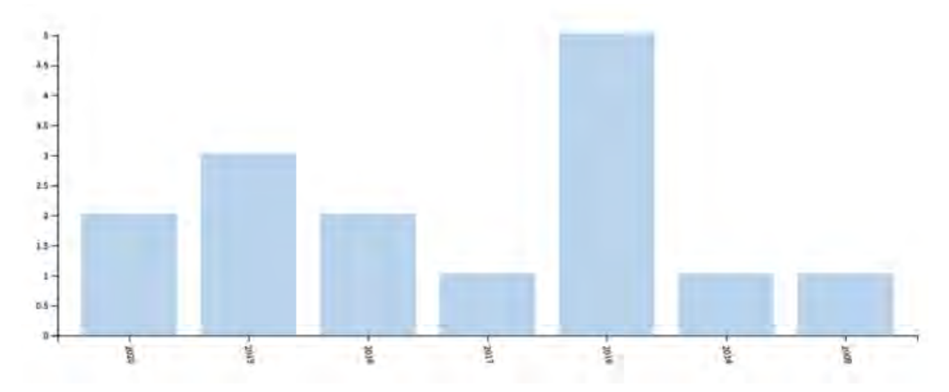
En esta gráfica, se aprecian tres crestas de máxima producción de artículos a saber: en 2016, 2018 y 2019. Ahora bien, realizando la búsqueda de la relación de cada una de las técnicas como palabras clave: **QCA, CSQCA, MVQCA** y **FSQCA** and **INNOVATION**, en el mismo periodo de 1984-2020 se tienen la **Gráficas: 3.3, 3.4, 3.5, 3.6, 3.7, 3.8 respectivamente.** Se destaca que no produjo resultados la combinación **MVQCA** con INNOVATION.

Gráfica 3.3. Publicaciones que relacionan al título QCA con INNOVATION por especialidad. Mapa de árbol: 15 documentos



Fuente: Web of Science Clarivate (2020).

Gráfica. 3.4. Publicaciones que relacionan al título QCA con INNOVATION por especialidad. Diagrama de barras: 15 documentos



Fuente: Web of Science Clarivate (2020).

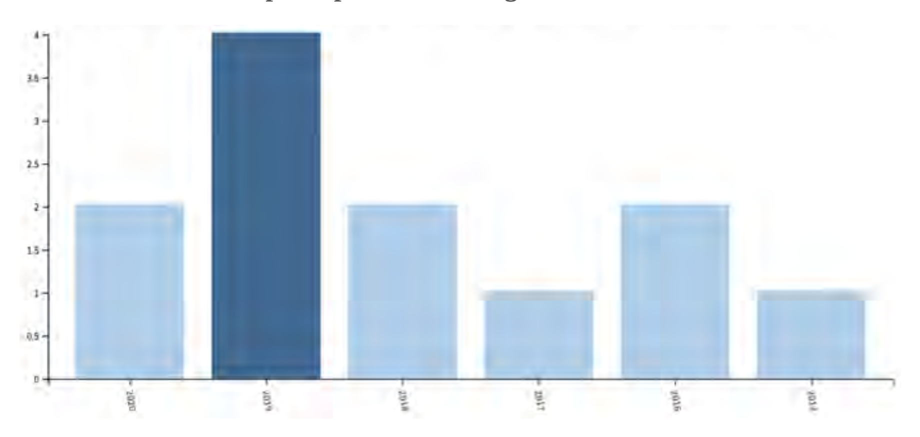
Gráfica. 3.5. Publicaciones que relacionan al título FSQCA con INNOVATION por especialidad. Mapa de árbol: 12 documentos



Fuente: Web of Science Clarivate (2020).

Se destacan: 8 documentos de negocios (*business*).

Gráfica. 3.6. Publicaciones que relacionan al título FSQCA con INNOVATION por especialidad. Diagrama de barras: 12 documentos



Fuente: Web of Science Clarivate (2020).

Producción modesta de 4 artículos relacionados en 2019. Se observa en el contexto general de **QCA**, que la producción de documentos apunta a tendencia de relacionarlo con técnicas estadísticas de carácter cuantitativo, métodos cualitativos y mixtos; así también a una alta tendencia de trabajar con la versión **fsQCA**; la relación con la innovación es aún incipiente.Como puede observarse, la primacía del uso único de **QCA** no evita que este método sea combinado con otros. En este proceso, la supremacía de los métodos estadísticos puede ser explicada tal vez por los conceptos de *complejidad causal o por los de necesidad y suficiencia*. Sin embargo,

> *"...la elección del método a combinar depende especialmente del carácter del fenómeno a ser estudiado, de su aplicación a este y de la evaluación de la factibilidad de sus resultados..."* **(Skaaning, 2006, p. 184).**

No hay que olvidar que el fin último de una investigación es su contribución al conocimiento científico.

Comparado con un diseño de investigación monométodo, la combinación de métodos es una estrategia de investigación de fuente de datos, recolección de datos, enfoques y de análisis de datos que permite estudiar un fenómeno determinado desde una perspectiva más amplia, integrando enfoques cualitativos y cuantitativos, con el fin de incrementar su comprensión y explicación **(Stange et al., 2006)**. Si bien es cierto que los *supuestos subyacentes a los métodos*, al ser combinados, pueden alterar los *supuestos ontológicos* de cada uno así como toda la investigación **(Fischer, 2011, p. 32)**,

> *"...la combinación de métodos sigue siendo desarrollada e implementada por investigadores en diversas disciplinas para incrementar la validez de los datos y los resultados".*

En la literatura sobre combinación de métodos, esta estrategia recibe el nombre de *métodos mixtos*, un diseño de investigación que comprende la *integración de técnicas cualitativas y cuantitativas* en diversas etapas que

van desde la formulación del problema, las preguntas de investigación, los objetivos, la recolección de datos, grado de análisis, análisis de datos y la integración de estos en un estudio único con una inferencia final **(Hesse-Biber, 2010)**.

De acuerdo a **Tashakkori y Teddlie (2003),** el diseño de investigación de *métodos mixtos* contiene dos distintas tipologías:

- *Diseño multi-método*: Este implica incluir más de un método pero una sola visión del mundo. Incluye estudios *cualitativos* de un lado y *cuantitativos*, por otro.
- *Diseño de método mixto*: Este implica usar métodos de recolección de datos *tanto cualitativos como cuantitativos* y métodos de investigación, en distintas fases de la investigación o en la fase exclusiva de métodos de investigación.

Razones para combinar QCA con otros métodos

La principal razón por combinar métodos en ciencia social orientada a la innovación, es para compensar aquellos problemas asociados al uso de un método único **(Fischer, 2011)**, como, por ejemplo, cuando un solo método no entrega todas las respuestas requeridas. Así, al aplicar múltiples perspectivas, se pueden obtener respuestas ya sean *explortorias y/o confirmatorias* desde distintas visiones del fenómeno.

Del mismo modo, la combinación de métodos permite llenar ciertos vacíos de información al permitir recolectarlos a través de métodos alternativos en pro de la variación y validez de los mismos y tratar datos en distintos niveles de análisis (macro, meso, micro) **(Morse, 2003, p. 205)**.

En el caso de **QCA**, unos de los aspectos más controvertidos que motiva a combinarlo con otros métodos *es el desafío de articularlo con estudios de caso* **(Ragin, 1987, p. 121; Berg-Schlosser, 2012, pp. 55-84)**. Para ello, se requiere de un *profundo conocimiento de los casos que debe generarse preferentemente antes de ejecutar el análisis* **(Rihoux y Álamos-Concha, 2013)**. El conflicto que surge aquí es entre dos objetivos:

- Por un lado, es oportuno *reunir evidencia de distintos casos y capturar la complejidad de los mismos* y,
- Por otro, *producir un cierto nivel de generalización* **(Rihoux y Álamos-Concha, 2013, p. 4; Ragin, 1987)**.

En consecuencia, las dos cuestiones que merecen ser abordadas es: hasta qué punto se puede integrar una investigación orientada al caso con **QCA** y cómo debiera ser realizado. **Rihoux y Álamos-Concha (2013),** abordan cinco características de una genuina investigación de *caso en profundidad*. Estos son:

1. Un estudio de caso en profundidad requiere recolectar mucha información e*n contextos complejos donde hay barreras idionuíticas y temporales.* En general, cuando el objeto de estudio se encuentra en el nivel meso-micro, la interacción con los individuos y las organizaciones son inminentes para recolectar los datos. Sin embargo, este proceso presenta diversas barreras y desafíos, tales como el idioma de las comunidades y zonas de un país; los distintos estilos de vida, temas de género y horarios; la censura que dificulta el acceso tanto a información valiosa como a informantes claves para la investigación así como conflictos y seguridad que dificulta la movilidad en distintas zonas de una región o país. Por lo tanto, realizar un *estudio de caso en profundidad* es bastante arduo bajo estas condiciones y más aún cuando el número de casos es elevado, porque cada caso en si merece un periodo determinado de estudio y campo. Por lo tanto, los resultados a obtener pueden llegar a ser poco satisfactorios.

2. El problema del sesgo en la información recolectada y las interpretaciones del investigador pueden ser potenciales o evidentes. El trabajo de campo o la utilización de métodos que nos acerquen al objeto de estudio con relativa *intimidad* son deseables. Si la información es extraída desde bases de datos previamente creadas, se corre el riesgo de tratar con información sesgada, y si se trabaja con *etnografía* entonces hay que evaluar el sesgo y la validez de los mismos. Tratar estas cosas antes de iniciar el

análisis **QCA** es imprescindible para realizar una correcta interpretación de los datos. Asimismo, es oportuno evaluar el grado de precisión de la evidencia en términos de fiabilidad cuando se recopila información de primera fuente.

3. El dilema entre el conocimiento en profundidad y el número de casos. Estudiar en profundidad dos casos puede llegar a ser manejable, no así con más de diez, más aún si el periodo de tiempo disponible para investigar es limitado. Si el objetivo es obtener cierta intimidad, los contextos en que la investigación se realiza generan ruido durante el proceso de investigación, pudiendo llevar a perder *cientificidad* del material obtenido, sobre todo cuando este no ha sido recolectado para fases post-investigación.
4. Falta de evidencia en la cadena de eventos: la *caja negra,* o fuerza causales que van desde las configuraciones al resultado de interés, no se estudia sistemáticamente, por lo tanto se requiere de explicación histórica para abordarlo. En consecuencia, existe una conexión parcial entre **QCA** y estudio de caso en profundidad. **QCA** se orienta a la comprensión del porqué de un fenómeno, mientras que para conocer lo que ocurre en la caja negra se requiere comprender cómo ocurren los hechos. En tanto que **QCA** se encuentra menos orientado a comprender los eventos subyacentes entre las configuraciones y el resultado de interés, se requiere de un análisis dentro del caso combinando métodos para este fin.
5. Desbalance en el conocimiento de cada caso. En un análisis comparativo se corre el riesgo de llegar a un *conocimiento desequilibrado* entre los casos, debido a las diversas *barreras contextuales*. Por ejemplo, si se realiza un trabajo de campo en una determinada zona con condiciones favorables para recolectar información, se puede lograr un buen nivel de profundidad y familiaridad con el caso, sin embargo, si la zona tiene problemas de seguridad, infraestructura, transporte y acceso, esta se hace más ardua para realizar entrevistas, observación o diarios. Este desbalance en la familiarización con cada caso afecta la validez de los resultados.

Los cinco desafíos expuestos demuestran que **QCA** no está completamente articulado con la investigación de estudios de caso. Para resolver estas limitaciones existen ciertas estrategias que hacen inminente combinar **QCA** con otros métodos. Esto forma parte de las buenas prácticas acerca de cómo tratar los principales desafíos de un estudio comparativo.

Formas de combinar con otros métodos: fase pre-QCA

Este proceso involucra distintas fases del diseño de investigación, desde la concepción del problema, la formulación de las preguntas de investigación, las fuentes de datos (*recolección*), los enfoques y el análisis de los datos. Además, a ello se suma su inclusión en una fase **post-QCA (Rihoux y Álamos-Concha, 2013)**, así que se tiene:

1. Preguntas de investigación bajo este enfoque. Por lo general, cuando se utiliza **QCA** las preguntas de investigación van orientadas a explicar el porqué de un fenómeno, es decir, responder a la pregunta de bajo qué combinación de factores un determinado fenómeno ocurre y no ocurre. Al combinar métodos, las preguntas pueden extenderse más allá de las configuraciones que explican un fenómeno, es decir, se pueden formular abordando el proceso bajo el cual los hechos ocurrieron. Por ejemplo, si se está interesado en los factores que motivan a implementar innovaciones bajo la nube (*cloud*) en una determinada empresa de servicios, una pregunta de investigación bajo este enfoque podría ser: primero, ¿por qué algunas empresas deciden hacerlo mediante la innovación por mercadotecnia digital y otras no? y, segundo, ¿cuáles son los mecanismos causales que motivan a una empresa de servicios a innovar a través del la nube (*cloud*) y a otras no? Ambas preguntas de investigación abordae dos cuestiones interrelacionadas: primero, conocer los factores y su combinación para explicar el porqué del fenómeno y, la segunda, permite abordar el fenómeno desde un enfoque orientado al proceso, el que debe ser estudiado en profundidad y con diversas técnicas de recolección de datos.

2. Recolección de datos en múltiples pasos. Esta estrategia permite obtener conocimiento en profundidad antes de ejecutar un análisis de **QCA**. La estrategia consiste en la búsqueda de múltiples fuentes cuantitativas y cualitativas:
   a. Entre las *cuantitativas* se pueden realizar encuestas, utilizar bases de datos elaboradas y disponibles por organizaciones nacionales e internacionales y entrevistas estructuradas.
   b. Por su parte, entre las *cualitativas* se pueden llevar a cabo observación etnográfica a corto-largo plazo, entrevistas semi-estructuradas o narrativas que permitan la interacción con informantes claves cuando la investigación es de nivel meso-micro, entre otros. Así, al lograr un conocimiento *íntimo* del caso, el investigador podrá realizar *interpretaciones desde una posición relativamente privilegiada en relación al nivel macro-meso de* **QCA (Rihoux y Lobe, 2009, p. 227)**.
3. Aplicación de triangulación para evitar sesgo y ganar precisión de los datos. En el proceso de recolección de datos, cuando, por ejemplo, se realiza una entrevista, existe una estrategia para evaluar esta información en términos de sesgo. La manera propicia es a través de la *observación*, es decir, se debe definir lo que el contenido de la observación es y evaluar el grado de precisión de la misma **(Beach y Pedersen, 2013, pp. 192-194)**. Aquí cabe preguntarse sobre el papel de los entrevistados durante la entrevista y decidir si la observación será tratada como primaria (directamente involucrada en el evento) o como material secundario. La precisión de los datos obtenidos debe ser mayor si se trata de una fuente primaria que de una secundaria **(Beach y Brun-Pedersen, 2013, p. 183)**. Sin embargo, muchas veces el entrevistado puede haber olvidado información clave sobre el problema de interés y reproducir una narrativa que no necesariamente es la propia sino aquella que ha escuchado por otros. Ello provoca problemas de precisión en la fuente primaria. Por otro lado, para manejar un bajo riesgo de sesgo, es oportuno aplicar una estrategia llamada *triangulación*, que significa aplicar distintas fuentes de recursos, por ejemplo, archivos, entrevistas, bases de datos, y mantener cierta independencia de los mismos **(Beach y**

Brun-Pedersen, 2013, p. 184). Esto último puede ser ejemplificado de la siguiente manera: al hacer entrevistas, se podría realizar un perfil del entrevistado con diversidad en edad del CEO, género, educación, lugar de residencia de la empresa, giro industrial, innovaciones practicadas, presupuesto de innovación, gastos por innovación entre otros.

4. Hacia una cantidad manejable de datos. Ganar *intimidad* con los casos es una tarea ardua y muchas veces el trabajo de campo o la aplicación de métodos tales como la etnografía y entrevistas no están al alcance de todos, ya sea por escasa financiación, seguridad, tiempo, idioma, etcétera, por lo que:

a. Una estrategia para adquirir información de dichos casos es mediante la *consulta a expertos, ya sean académicos, consultores, periodistas especializados,* entre otros; pero es necesario que en el momento de ejecutar el análisis **QCA**, el investigador esté familiarizado con cada caso **(Rihoux y Lobe, 2009, p. 231**).

b. Otra estrategia es crear un puente entre **QCA** y casos de estudios etnográficos para ganar conocimiento en profundidad si la ***N*** *es pequeña* (**Rihoux y Lobe, 2009, p. 239**). Para **N** *grande* aún no existe consenso sobre la mejor manera de obtener familiaridad con los casos, aunque muchos investigadores reconocen que **QCA** es adecuado para diseño con **N** *grande* y que también es posible ganar conocimiento de diferentes categorías de casos **(Schneider y Wagemann, 2012; Rihoux, et al., 2013)**.

c. Una tercera estrategia para enfrentar estas barreras contextuales puede ser la integración en proyectos de investigación o la creación de redes de determinados expertos, permitiendo un trabajo colectivo e interdisciplinario para crear conocimiento en distintas fases de un proyecto **(Rihoux y Lobe, 2009, p. 239)**.

Formas de combinar con otros métodos: fase post- QCA

El objetivo de las fases **pre-QCA, durante QCA** y **post-QCA**, en su combinación con la investigación de *estudios de caso,* es articular ambos de manera real. Como se ha abordado, el diálogo con los casos individuales se lleva a cabo primeramente durante el proceso de *recolección de datos*, el cual está bastante distante del trabajo que **QCA** realiza después.

En el proceso de análisis de los datos, el siguiente paso es interpretar los diferentes caminos obtenidos a través de las distintas soluciones que se generan: *compleja, parsimoniosa e intermedia* **(Rihoux y Lobe, 2009, pp. 235-237)**. Esta etapa permite comprender la *selección de los casos* y, a través *de la vuelta* a los mismos, se construyen las narrativas de cada uno de ellos para movemos hacia la complejidad **(Curchod et al., 2004)** y se plantean tres perspectivas desde las cuales se pueden *interpretar las soluciones:*

1. *Interpretación caso a caso* mediante el uso de los factores explicativos claves indicados en las soluciones traduciendo los factores de la fórmula en narrativa causal **(Rihoux y Lobe, 2009, p. 236)**.
2. *Interpretación entre los casos*, donde se identifican similitudes, diferencias y patrones entre las distintas nanativas, construidas a partir de las soluciones **(Rihoux y Lobe, 2010, p. 236)**.
3. *Generalización histórica limitada,* donde se va más allá de los casos observados con el fin de apoyar la generalización histórica limitada **(Ragin, 1987, p. 31)**. En particular se pueden formular proposiciones no solo de una comparación sistemática de los casos sino hacia otros casos similares que comparten un número determinado de características en el proceso de **QCA**.

Hasta aquí se ha argumentado sobre la importancia de articular **QCA** con casos de estudios y cómo el conocimiento basado en el caso juega un papel decisivo en el protocolo de **QCA**. Según lo abordado hasta el momento, parece que **QCA** está más bien orientado a la configuración que al caso,

puesto que este comprende la causalidad en términos de *configuración* de factores explicativos que contribuyen a un resultado determinado **(Blatter y Haverland, 2012)**. En este sentido, un enfoque orientado a la combinación de métodos es adecuado para articular análisis de casos de estudio en profundidad y **QCA** (Rihoux y Ragin, 2009) en una fase **post-QCA**. Algunos de los casos que se pueden tratar después de ejecutar un análisis de **QCA** es *uno típico, uno desviado y uno contradictorio* **(Blatter y Haverland , 2012, p. 231; Schneider y Wagemann, 2012, pp. 305-312)**. Para efectuar los análisis de estos casos se pueden aplicar ciertas técnicas orientadas al proceso, tales como: *análisis histórico, técnicas* de *Process-Tracing* **(Blatter y Haverland, 2012, pp. 95-97)**, *análisis de congruencia*, entre otros. Así, combinar **QCA** con otros métodos logra el objetivo de conocer las causas y los mecanismos detrás de un determinado fenómeno.

A continuación, en la **Tabla 3.26**, se ilustran once principios para la selección de casos.

Tabla 3.26. Resumen de los principios de selecciónde casos

	csQCA y fsQCA	fsQCA
Caso único y Process-Tracing comparativo	*Principios de selección de casos diversos*: escoger al menos un caso para cada vía de la solución	*Principio de máxima pertenencia en el conjunto*: el caso más típico manifiesta máxima puntuación de pertenencia de conjunto en el subconjunto y el superconjunto
	Principio de pertenencia única; escoger casos que son solo miembros de una vía	*Principio de máxima diferencia de pertenencia en el conjunto*: B caso más desviado manifiesta máxíma diferencia en su membresía de conjunto en el subconjunto y el superconjunto
	Principio de suficiencia de la Tabla de Verdad: para escoger un caso desviado por cobertura, determinar la fila de la Tabla de Verdad del caso	*Principio de diferencia en tipo y grado*: diferencias en grado deben solo ser establecidas entre casos que son similares entipo y localizados en el mismo lado de la diagonal secundarla
Process-Tracing comparativo	*Principio del resultado positivo:* Al menos un caso debe ser un miembro del resultado en PT comparativo	*Principio de desviación en tipo: escoger* casos desviados por consistencia que son cualitativamente diferentes de los casos típicos en su membresía en el superconjunto
	Principio de necesidad de la Tabla de Verdad: cuando se compara un caso típico y un caso individual Irrelevante, se escogen dos casos que difieren en su membresía en el factor necesario y el resultado, pero que comparten la membresía cualitativa en todos los otros factores que constituyen la Tabla de Verdad	*Principio de máx-máx diferencia:* Cuando se comparan dos casos tfpicos o un caso típico con un caso Individualmente
		Principio de máx-min diferencia: cuando se compara un caso típico con un caso desviado, se maximiza la diferencia de la membresía del conjunto de casos en el superconjunto y se minimiza la diferencia en el subconjunto

Fuente: Schneider y Wagemann (2012, p. 311).

Estos principios pueden ser estudiados con *Process-Tracing* **(PT)** en una fase **post-OCA.** Algunos de estos principios son generales porque son aplicables a dos de las variantes del **QCA** y a las *relaciones de necesidad y suficiencia*, así como a **PT** de casos únicos y comparativos. Por otro lado, otros son aplicables específicamente a estudios de casos comparativos. Por último, hay cinco principios que son específicos al **fsQCA**. En términos teóricos, los once principios *no son mutuamente excluyentes*, sin embargo, en la práctica el investigador puede verse enfrentado a la escasez de tiempo y a datos incompletos, por lo que la selección de alguno de estos principios será menester **(Schneider y Wagemann, 2012, p. 311)**. Estas estralegias para la selección de casos de la fase **post-QCA** ayudan a mejorar *la calidad de la relación de teoría-conjunto y a promover la generación de inferencias causales* **(Schneider y Rohlfing, 2013, p. 5)**.

La combinación de métodos es una estrategia de investigación justificable cuando el fenómeno a investigar adquiere resultados que son incompletos en su explicación, cuando con más de un método/ enfoque/ fuentes a aplicarse puede llegar a respuestas fiables y exhaustivas. El predominio del uso de **QCA** como único método de investigación no descarta la importancia de combinarlo con métodos cualitativos, cuantitativos y ambos, puesto que el desafío no solo es ampliar la visión desde la cual se estudia un fenómeno y completar respuestas, sino, además, lograr el desafío de articular **QCA** con casos de estudio. Se ha visto que este desafío abarca tanto las fases de **pre-QCA** como de **post-QCA**, tanto en la recogida de datos como en el análisis de los mismos. Se ha señalado además que existen técnicas de investigación como el *Process-Tracing* que permiten ir más allá del *porqué* del fenómeno, es decir, no solo orientar el estudio a las configuraciones que contribuyen a un determinado resultado, sino además, fijarnos en el *cómo* a través de la búsqueda y observación de los mecanismos causales que producen un resultado. Esta combinación permite enriquecer la investigaeión y presentar respuestas más robustas sobre las causas de un fenómeno determinado, su comprensión y su explicación.

Casos csQCA

A continuación, se presentarán una serie de ejercicios con el software **fsQCA 3.0**, para que lo conozca y comprenda los conceptos hasta aquí vertidos.

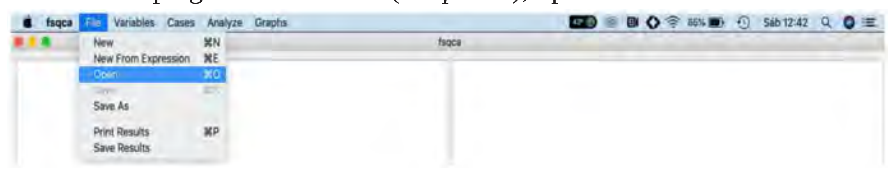
Manejo de conjunto de datos nítidos (*crisp-sets*) en csQCA. Caso 1

Como ejemplo se tomará el caso de las herramientas de innovación por mercadotecnia digital **Mejía-Trejo (2017c)**, y que describe la táctica empleada por 18 empresas que la practican, la cual se encuentra en el archivo de trabajo: **tácticas de mercadotecnia digital.cvs** y que se compone de los siguientes elementos, de los que se evaluará su dicotomía:

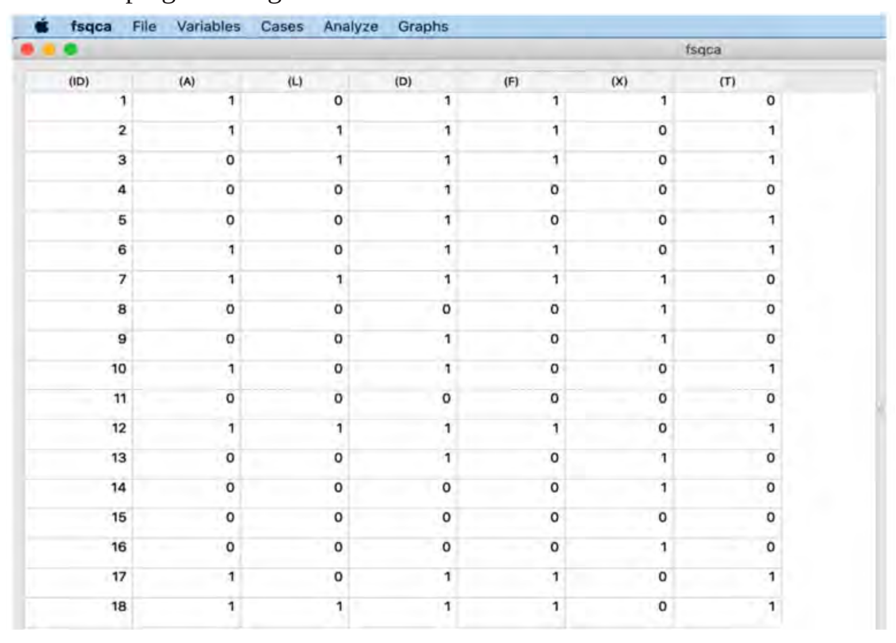
a. Acciones que generan conciencia, de la presencia de la empresa y sus productos / servicios, en el consumidor de internet**, A.**
b. Acciones que generan compromiso y lealtad a la empresa y sus productos/ servicios, en el consumidor de internet**, L.**
c. Acciones que generan deseo y experiencia la empresa y sus productos / servicios, en el consumidor de internet**, D.**
d. Acciones que generan efectividad en el call to action de la empresa y sus productos / servicios, en el consumidor de internet**, F.**
e. Acceso del consumidor de internet a los medios digitales**, X.**
f. Efectividad de las tácticas por innovación en mercadotecnia digital**, T.**
g. Empresa, **ID**.

Con el archivo de trabajo: **tácticas de mercadotecnia digital.cvs,** abra y despliegue los datos en el programa **fsQCA,** como se solicita a continuación:

1. Para desplegar datos nítidos (*crisp-sets*), oprima:



2. Se desplegará la siguiente ventana:



(ID)	(A)	(L)	(D)	(F)	(X)	(T)
1	1	0	1	1	1	0
2	1	1	1	1	0	1
3	0	1	1	1	0	1
4	0	0	1	0	0	0
5	0	0	1	0	0	1
6	1	0	1	1	0	1
7	1	1	1	1	1	0
8	0	0	0	0	1	0
9	0	0	1	0	1	0
10	1	0	1	0	0	1
11	0	0	0	0	0	0
12	1	1	1	1	0	1
13	0	0	1	0	1	0
14	0	0	0	0	1	0
15	0	0	0	0	0	0
16	0	0	0	0	1	0
17	1	0	1	1	0	1
18	1	1	1	1	0	1

La versión del software **fsQCA 3.0** (julio de 2017) contiene un método para realizar análisis de datos nítidos: el algoritmo de *tabla de verdad*, método que hace uso del *algoritmo* **Quine-McCluskey**. Sobre el caso del archivo de trabajo: **tácticas de mercadotecnia digital.cvs.**

Tabla de verdad.
Caso 2

Dos tareas importantes, estructuran la aplicación de la *tabla de verdad* en el conjunto de datos nítidos:

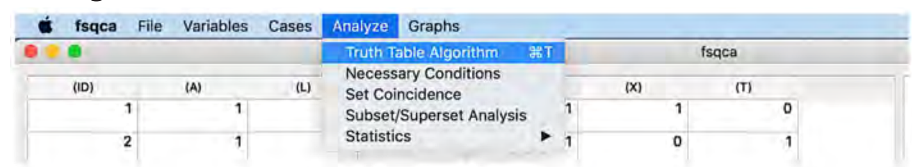
- La evaluación de la distribución de casos, entre diferentes combinaciones lógicamente posibles de las *condiciones causales*, y
- La evaluación de la *consistencia* de la evidencia para cada combinación causal, con el argumento de que los casos, con esta combinación de condiciones, constituyen un subconjunto de los casos con el resultado. Es decir, comparten el resultado en cuestión.

La *tabla de verdad*, implica un procedimiento analítico de dos pasos:

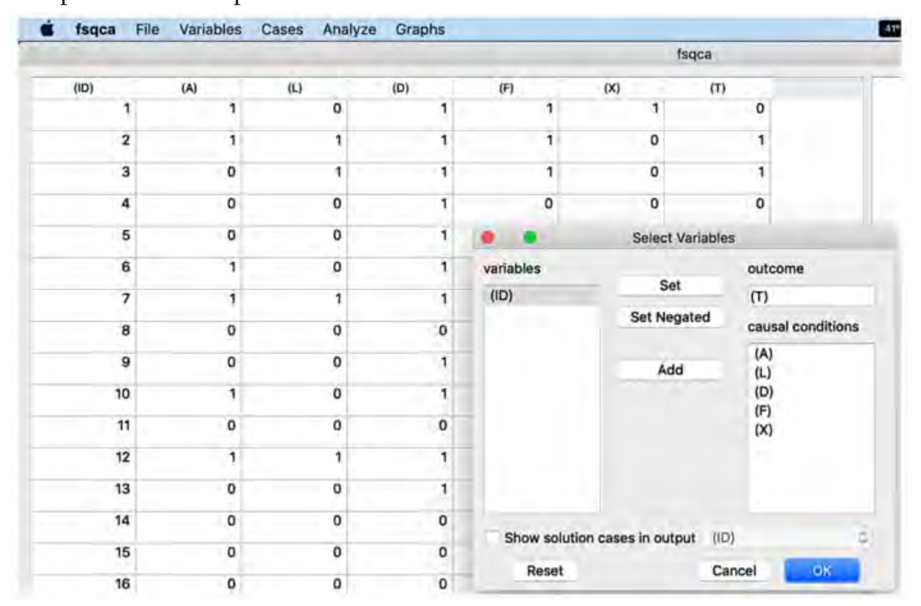
- El primer paso, consiste en crear una hoja de cálculo de la *tabla de verdad* a partir de los datos sin procesar, que implica principalmente especificar el resultado y las *condiciones causales* para incluir en el análisis.
- El segundo paso, consiste en preparar la hoja de cálculo de la *tabla de verdad* para el análisis, seleccionando un umbral de frecuencia y un umbral de *consistencia*.

Así, del archivo de trabajo: **tácticas de mercadotecnia digital.cvs.,** realice:

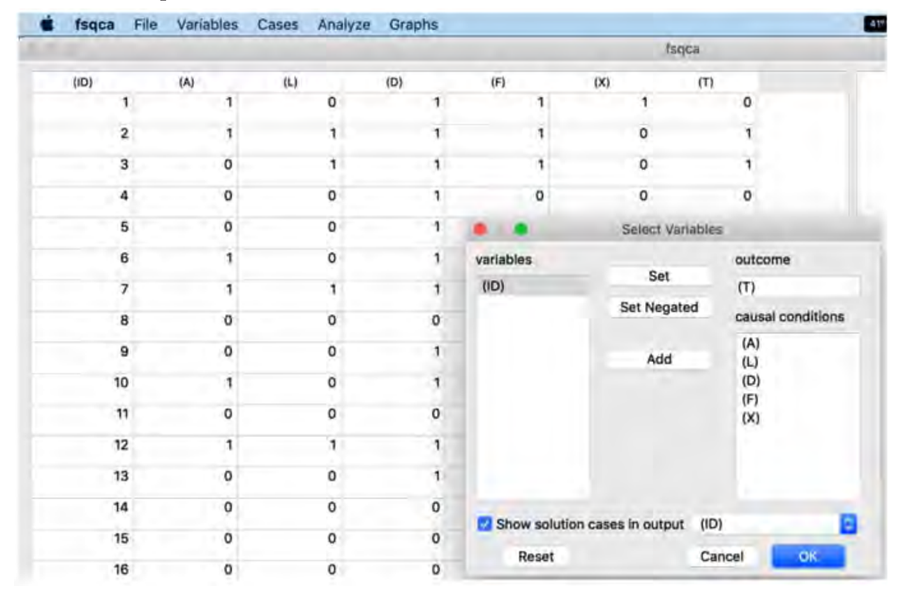
1. Para crear la *tabla de verdad*, oprima: **Analyze → Truth Table Algorithm…**



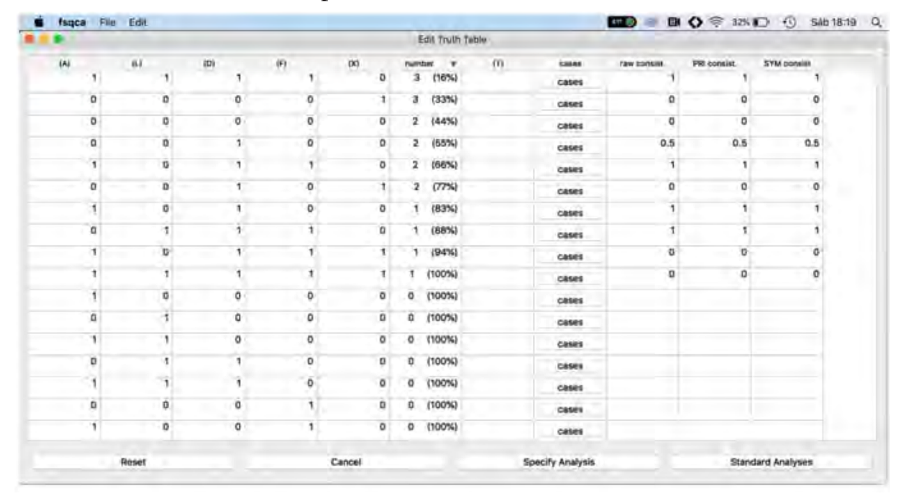
2. Agregar las variables según se requieran… Identifique y resalte el aspecto del caso que desea explicar y transfiéralo al campo **Outcome** oprimiendo **Set** …Seleccione de la lista preliminar, las *condiciones causales*, resaltando una a la vez y oprimiendo **Add** para moverlas una por una al campo **Causal Conditions**…



3. Marque la casilla junto a **Show solution cases in output** y elija la variable que sea su **ID** de caso.



4. Oprima en el botón **OK** y aparecerá la siguiente ventana, que contiene la *tabla de verdad* completa:



5. La *tabla de verdad* tendrá $\mathbf{2^k}$ filas (donde ***k*** representa el número de *condiciones causales*), mostrando todas las combinaciones posibles de las mismas (desplácese hacia abajo para ver todas las combinaciones posibles).

Incluso, se muestran las configuraciones de condiciones para los que no existen casos empíricos, llamadas *remanentes o residuales* **(Rihoux y De Meur, 2009, p. 59)**. En la columna (**T**), se debería indicar si una cierta configuración debe ser incluida para la minimización, pero en esta etapa está todavía vacía puesto que no se ha decidido aún esta situación. Los **1s** y **0s**, representan tanto a la pertenencia completa como a la no pertenencia completa cero, para cada condición y de forma respectiva. Para cada fila, se crea un valor con la siguiente descripción en la **Tabla 3.27.**

En nuestro ejemplo, las filas: uno, cinco, siete y ocho, tienen una *raw consistency* (*consistencia bruta*) de **(1),** lo cual significa que todos los casos que tenían estas configuraciones de condiciones explicativas compartían el resultado de interés **(T= 1),** mientras las filas: dos, tres, seis, nueve y diez encontramos una configuración de condiciones cuyo único caso no tenía el *resultado de interés* **(Y=0),** *consistencia bruta de* **(0).**

La configuración de condiciones más interesante es la de la fila cuatro, para la cual existen dos casos, y en la que la consistencia bruta es de **(0.5).** Esto significa que de estos dos casos, solo uno tenía el resultado de interés. Este tipo de configuración de condiciones se denomina *contradicción*, porque en ella la misma configuración de condiciones explicativas conduce a resultados contrarios **(Rihoux y De Meur, 2009, p. 48)**. Las *contradicciones* en análisis **QCA** *deben ser evitadas en la medida de lo posible.* Es frecuente que estas *contradicciones* apunten a una *condición explicativa* oculta que originalmente no fue tomada en cuenta en la investigación, y que debería explicar la divergencia en cuanto a los resultados.

Tabla 3.27. Descripción de valores que reporta fsQCA en su versión conjunto de datos nítidos (crisp-sets)

Valor	Descripción
number	Es el número de casos, que muestran la combinación de condiciones, es decir, el número de casos existentes para cada configuración de condiciones.
raw consist.	Es la proporción de casos, en cada fila de la *tabla de verdad*, que muestran el resultado. Para comprender cuántos casos con el *resultado de interés* están cubiertos por cada configuración de condiciones, se utiliza este término como *índice de consistencia bruta*. Este índice indica qué porcentaje de casos dentro de una configuración de condiciones muestra el resultado de interés entre el número total de casos en dicha configuración. La *consistencia bruta* es importante ya que antes de realizar la minimización booleana debemos indicar cuáles de las configuraciones serán incluidas en la operación, y en particular debemos seleccionar el número de casos que las configuraciones deben tener para ser incluidos. La opción por default es (**1**) y los niveles de consistencia mínimos de las configuraciones que se incluirán. El nivel de consistencia mínimo aceptado rondaría **0.8** o incluso **0.85**, aunque niveles más altos de consistencia conducen a resultados más robustos, debiendo siempre ser explicitada y justificada la elección hecha a este respecto (véase para una discusión, **Ragin, 2004**). En particular, en **csQCA** es preferible incluir solo aquellas configuraciones con una consistencia de (**1**) pues niveles más bajos indican la existencia de una *contradicción*. En análisis de conjuntos difusos la interpretación del índice de consistencia bruta es más compleja, ya que en este tipo de análisis los casos tienen una inclusión difusa en cada una de las condiciones de la configuración **(Ragin, 2009)**. En el capítulo sobre, **fsQCA** se explica cómo se calcula e interpretan las relaciones de *suficiencia*. En caso de duda sobre el valor de *consistencia*, tanto para **csQCA** como para **fsQCA**, se habilitan dos parámetros complementarios: **PRI** y **SYM,** analizados en detalle por **Schneider y Wagernann (2012, pp. 242-243)**.
PRI consist.	Llamado *Propotional Reduction in Inconsistency,* es la medida alternativa de *consistencia* (desarrollada para *datos difusos*), basada en una reducción cuasi proporcional en el cálculo de errores. En análisis de datos nítidos, esto será igual a *raw-consist.* El **PRI**, que no puede entenderse como sustituto del valor de *consistencia*, indica en qué medida una determinada configuración forma parte de (**Y**) y no de (**-Y**) o, en terminología de conjuntos, en qué medida (**X**) es un subconjunto de (**Y**) y no de (**-Y**) **(Schneider y Wagemann, 2012, p. 242)**. Siguiendo a los autores, se busca que las configuraciones sean consistentes con la idea de subconjuntos; así, un valor **PRI** bajo estará indicando que los valores de consistencia de (**X**) para (**Y**) e (**-Y**) están muy cerca, por lo que (**X**) fomará parte de ambos conjuntos. De lo contrario, se buscan valores altos en el parámetro **PRI**.
SYM consist.	Es la medida alternativa de *consistencia* para *datos difusos* basada en una versión simétrica de **PRI.**

Para tratar con estas *contradicciones*, existe una serie de estrategias a seguir que deben ser justificadas explícitamente y que pueden ser aplicadas antes y durante el proceso de minimización **(Ragin, 1987, pp. 113-118; Rihoux y De Meur, 2009, pp. 48-49; Schneider y Wagemann, 2012, pp. 120-123)** así también es indispensable considerar las medidas de *consistencia*, como una estrategia complementaria **(Schneider y Wagemann, 2012, p. 123)**. Por lo tanto, se recomienda:

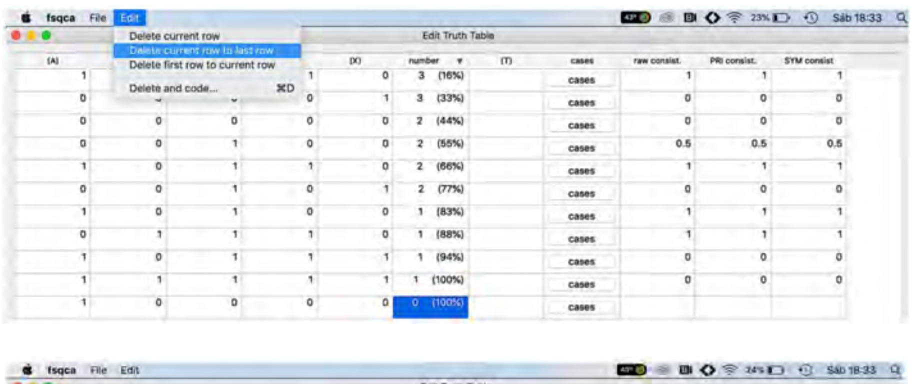
- *Agregar otra condición al modelo*. Esta estrategia permite una mayor diferenciación entre los casos, una mayor complejidad del modelo, y una aún posible *parsimonia*, lo que contribuye al *impedimento de las contradicciones*. Este paso debe ser debidamente justificado ya sea teórica o empíricamente, para evitar caer en una diversidad limitada mayor **(Rihoux y De Meur, 2009, p. 48)**.
- *Sacar condiciones del modelo y reemplazarlas por otras*.
- *Reexaminar la operacionalización de las condiciones* .
- *Reexaminar la definición y operacionalización del resultado de interés*.
- *Reexaminar los casos contenidos en las configuraciones contradictorias y evaluar la diversidad entre ellos.*
- *Reexaminar si los casos bajo estudio son parte de la misma población*. Este proceso puede llevar a *incluir/excluir* ciertos casos siempre y cuando exista una justificación teórica.
- *Recodificar todas las configuraciones contradictorias con valor (***0***) en la columna del resultado*, lo cual implica excluirlas del proceso de minimización.
- *Incluir todas las filas contradictorias en el análisis* **(Schneider y Wagemann, 2012, p. 122)**, lo cual se justifica por su contribución a la ocurrencia del resultado. La ventaja es que la solución obtenida cubrirá todos los casos estudiados y las vías obtenidas serán conjunciones de condiciones; sin embargo, su desventaja es que algunos casos que no son miembros del resultado serán cubiertos por la solución (*ibídem*).

- *Aplicar el criterio de orientación del resultado*, lo cual indica que la vía más frecuente para llegar al resultado en un número mayor de casos es el considerado.

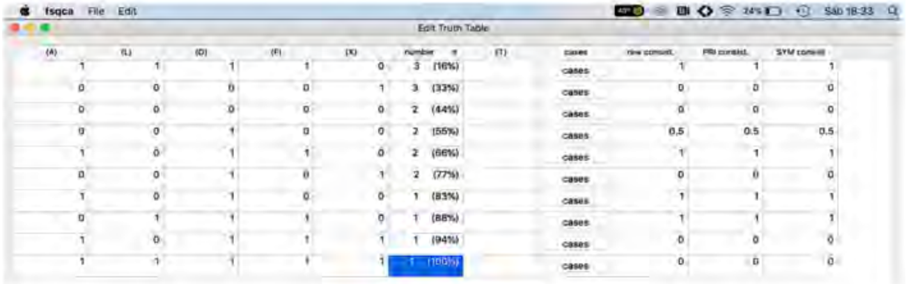
Ajustando el resultado (*Outcome*). Caso 3

Tenga en cuenta que la columna etiquetada como el **Outcome** (Campo **T**, en este ejemplo,) está en blanco. Depende del investigador, determinar el resultado para cada configuración utilizando el siguiente procedimiento:

1. El investigador, debe comenzar desarrollando una regla para clasificar algunas combinaciones (filas), como *relevantes* y otras como *irrelevantes*, en función de su frecuencia. Esto se logra seleccionando un umbral de frecuencia basado en el número de casos en cada fila, que se muestra en la columna **number**. Cuando el número total de casos **N**, en un análisis es relativamente pequeño, el umbral de frecuencia debe ser **1** o **2.** Cuando **N** es grande, sin embargo, se debe usar un umbral más sustancial. Es muy importante examinar la distribución de casos, entre las combinaciones causales.
2. Las *configuraciones* (filas), se pueden ordenar por su frecuencia (descendente o ascendente), oprimiendo en el encabezado de la columna **number**.
3. Después de ordenar las filas y seleccionar un umbral de frecuencia, elimine todas las filas que no cumplan el umbral. Si los casos se han ordenado en orden descendente según la columna number, oprima en el primer caso que esté por debajo del umbral y luego seleccione: **Edit → Delete current row last row.**



(A)	[illegible]	[illegible]		(X)	number	(T)	cases	raw consist.	PRI consist.	SYM consist.
1	[illegible]	[illegible]	1	0	3 (16%)		cases	1	1	1
0	[illegible]	[illegible]	0	1	3 (33%)		cases	0	0	0
0	0	0	0	0	2 (44%)		cases	0	0	0
0	0	1	0	0	2 (55%)		cases	0.5	0.5	0.5
1	0	1	1	0	2 (66%)		cases	1	1	1
0	0	1	0	1	2 (77%)		cases	0	0	0
1	0	1	0	0	1 (83%)		cases	1	1	1
0	1	1	1	0	1 (88%)		cases	1	1	1
1	0	1	1	1	1 (94%)		cases	0	0	0
1	1	1	1	1	1 (100%)		cases	0	0	0
1	0	0	0	0	0 (100%)		cases			



(A)	(L)	(D)	(F)	(X)	number	(T)	cases	raw consist.	PRI consist.	SYM consist.
1	1	1	1	0	3 (16%)		cases	1	1	1
0	0	0	0	1	3 (33%)		cases	0	0	0
0	0	0	0	0	2 (44%)		cases	0	0	0
0	0	1	0	0	2 (55%)		cases	0.5	0.5	0.5
1	0	1	1	0	2 (66%)		cases	1	1	1
0	0	1	0	1	2 (77%)		cases	0	0	0
1	0	1	0	0	1 (83%)		cases	1	1	1
0	1	1	1	0	1 (88%)		cases	1	1	1
1	0	1	1	1	1 (94%)		cases	0	0	0
1	1	1	1	1	1 (100%)		cases	0	0	0

Si los casos no se han ordenado, entonces, aquellos casos que no cumplan con el umbral se pueden eliminar individualmente seleccionando la fila que elija: **Edit → Delete current now.**

4. El siguiente paso, es distinguir las *configuraciones* que son subconjuntos consistentes del resultado, de las que no lo son. Para el conjunto de datos nítidos, esta determinación se realiza utilizando la medida de *consistencia* del marco teórico. La cual se reporta en la columna *raw consist*. Los valores por debajo de **0.75** indican una *inconsistencia* sustancial. Es útil ordenar los puntajes de *consistencia* en orden descendente, para evaluar su distribución (esto debe hacerse después de eliminar las filas que no alcanzan el umbral de frecuencia). La ordenación se realiza oprimiendo la cabecera de columna *raw consist.*

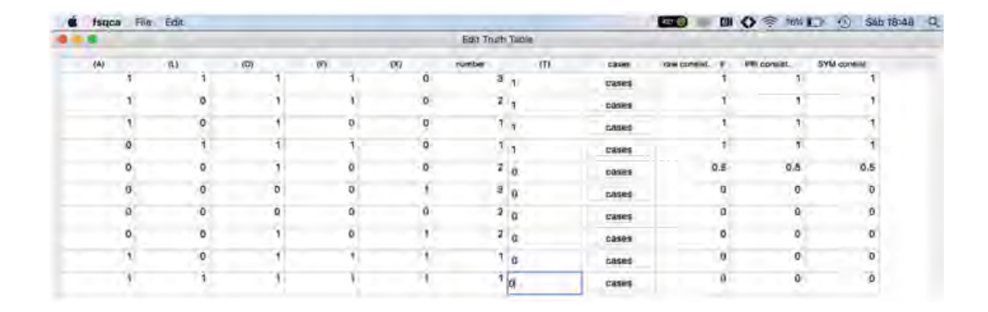


(A)	(L)	(D)	(F)	(K)	number	(T)	cases	raw consist.	PRI consist.	SYM consist
0	0	0	0	1	3		cases	0	0	0
0	0	0	0	0	2		cases	0	0	0
0	0	1	0	1	2		cases	0	0	0
1	0	1	1	1	[illegible]		cases	0	0	0
1	1	1	1	1	1		cases	0	0	0
0	0	1	0	0	2		cases	0.5	0.5	0.5
1	1	1	1	0	3		cases	1	1	1
1	0	1	1	0	2		cases	1	1	1
1	0	1	0	0	1		cases	1	1	1
0	1	1	1	0	1		cases	1	1	1



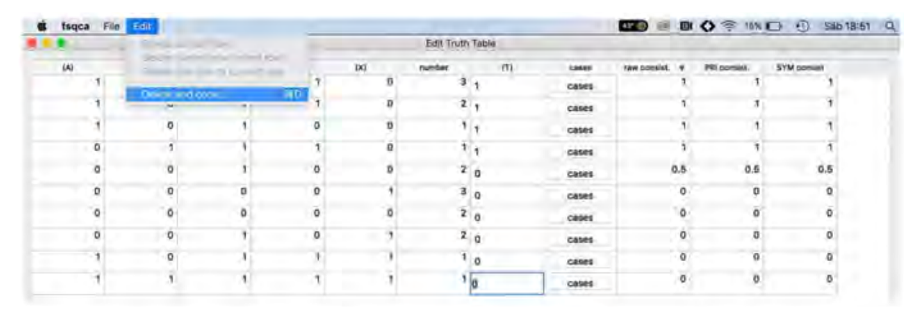
(A)	(L)	(D)	(F)	(K)	number	(T)	cases	raw consist.	PRI consist.	SYM consist
1	1	1	1	0	[illegible]		cases	1	1	1
1	0	1	1	0	2		cases	1	1	1
1	0	1	0	0	1		cases	1	1	1
0	1	1	1	0	1		cases	1	1	1
0	0	1	0	0	[illegible]		cases	0.5	0.5	0.5
0	0	0	0	1	[illegible]		cases	0	0	0
0	0	0	0	0	3		cases	0	0	0
0	0	1	0	1	2		cases	0	0	0
1	0	1	1	1	1		cases	0	0	0
1	1	1	1	1	1		cases	0	0	0

5. Identifique cualquier brecha, en el rango superior de *consistencia* que pueda ser útil para establecer un umbral de *consistencia*. Tenga en cuenta, que siempre es posible examinar varios umbrales diferentes y evaluar las consecuencias de reducir y/o aumentar, el límite de *consistencia*.
6. Ahora, es necesario indicar qué *configuraciones* pueden considerarse subconjuntos del resultado y cuáles no (ver también, el método alternativo a continuación). Ingrese un **(1),** en la columna de resultados (campo **T**, en este ejemplo) para cada configuración, cuyo nivel de *consistencia* cumpla o supere el umbral. Ingrese un **(0)**, en la columna de resultados para cada configuración cuyo nivel de *consistencia* no cumpla con el umbral de *consistencia*.

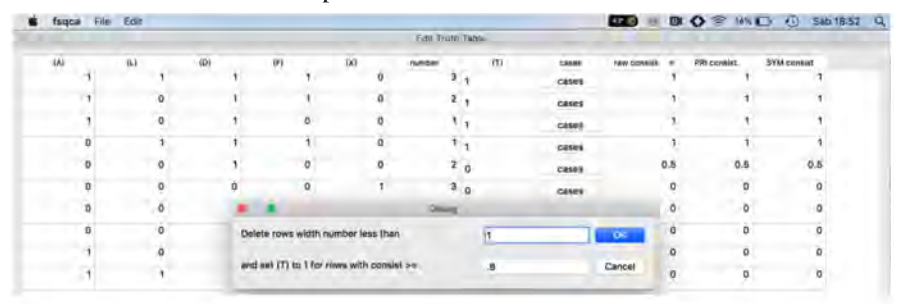


(A)	(L)	(D)	(F)	(X)	number	(T)	cases	raw consist.	PRI consist.	SYM consist
1	1	1	1	0	3	1	cases	1	1	1
1	0	1	1	0	2	1	cases	1	1	1
1	0	1	0	0	1	1	cases	1	1	1
0	1	1	1	0	1	1	cases	1	1	1
0	0	1	0	0	2	0	cases	0.5	0.5	0.5
0	0	0	0	1	3	0	cases	0	0	0
0	0	0	0	0	2	0	cases	0	0	0
0	0	1	0	1	2	0	cases	0	0	0
1	0	1	1	1	1	0	cases	0	0	0
1	1	1	1	1	1	0	cases	0	0	0

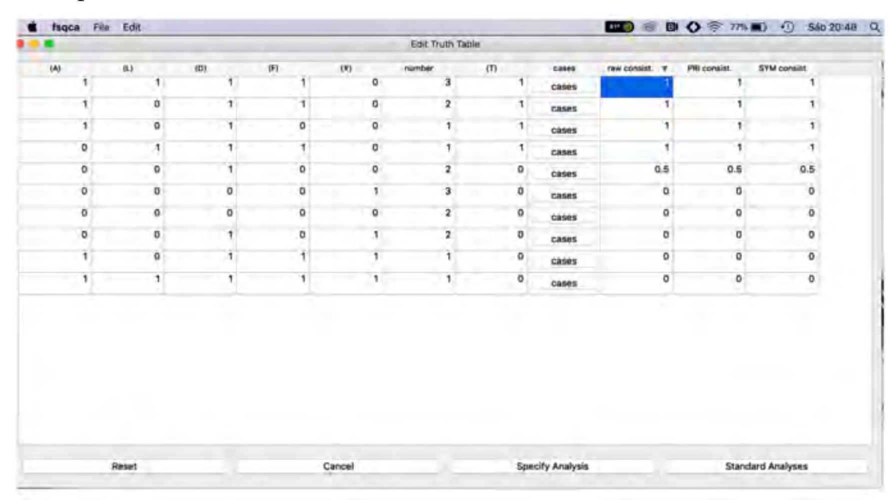
7. Alternativamente, uno puede usar la función **Delete and code** para automatizar este proceso. Seleccione: **Edit → Delete and code.**



En el primer campo, se selecciona el umbral de frecuencia. El número predeterminado de casos es **(1)**, pero puede cambiarse escribiendo el umbral de frecuencia seleccionado en el campo. En el segundo campo, se selecciona el umbral de *consistencia*. La *consistencia* predeterminada es **0.8**, pero esto puede cambiarse escribiendo el umbral de *consistencia* seleccionado en el campo.



Oprima el botón **OK**. El programa eliminará las filas donde no se alcanza el umbral de frecuencia y codificará el resultado como **(0)** o **(1)** dependiendo del umbral de *consistencia* seleccionado.



(A)	(L)	(D)	(F)	[illegible]	number	(T)	cases	raw consist.	PRI consist.	SYM consist.
1	1	1	1	0	3	1	cases	1	1	1
1	0	1	1	0	2	1	cases	1	1	1
1	0	1	0	0	1	1	cases	1	1	1
0	1	1	1	0	1	1	cases	1	1	1
0	0	1	0	0	2	0	cases	0.5	0.5	0.5
0	0	0	0	1	3	0	cases	0	0	0
0	0	0	0	0	2	0	cases	0	0	0
0	0	1	0	1	2	0	cases	0	0	0
1	0	1	1	1	1	0	cases	0	0	0
1	1	1	1	1	1	0	cases	0	0	0

8. La ventana, muestra la *tabla de verdad* que aparece después de:

a. Aplicar un umbral de frecuencia de **(1)**, a los datos y eliminando las *configuraciones* que no tienen observaciones (**6** *configuraciones*).

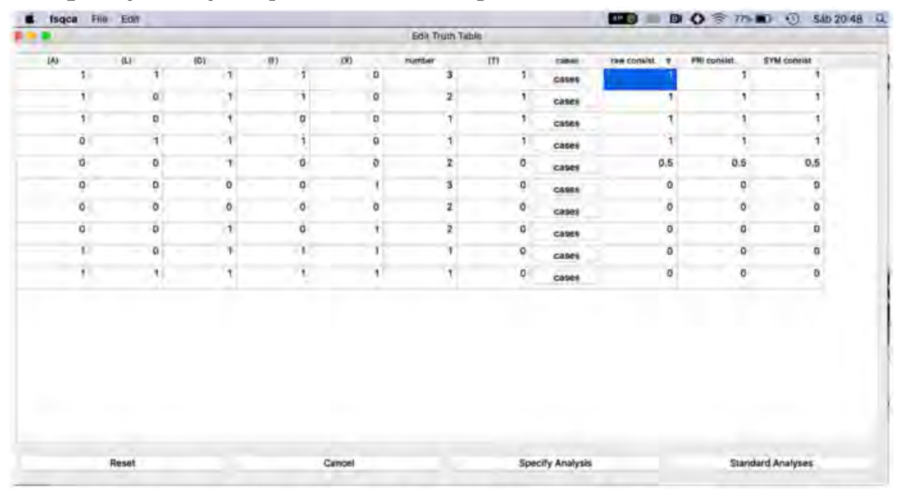
b. Seleccionando un umbral de *consistencia* de **0.9** y colocando un **(1)** en la columna **T**, para *con iguraciones* con *consistencia* **de 0.90** o mayor (**4** *configuraciones*) y un **(0)** para casos con *consistencia* más baja (**6** *configuraciones*).

Instrucción: *Specify Analysis Option.*
Caso 4

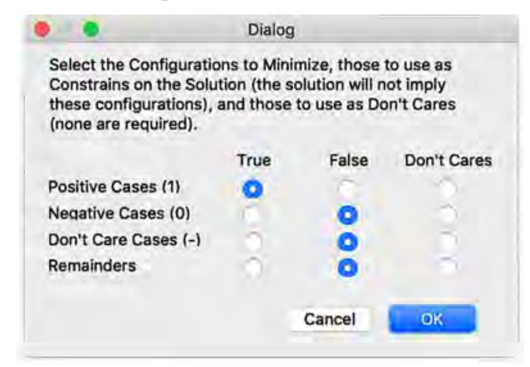
A partir de aquí, hay dos posibilidades para el análisis: especificar un análisis único (**Specify Analyses Option**) vs. generar los tres análisis estándar (**Standard Analyses** *complejo, parsimonioso e intermedio*). El procedimiento recomendado a seguir, es oprimir el botón **Standard Analyses** (de la que se recomienda, aplicar la generación de solución intermedia) y explicando a detalle, el funcionamiento de ambas opciones.

Para lograrlo, se debe realizar:

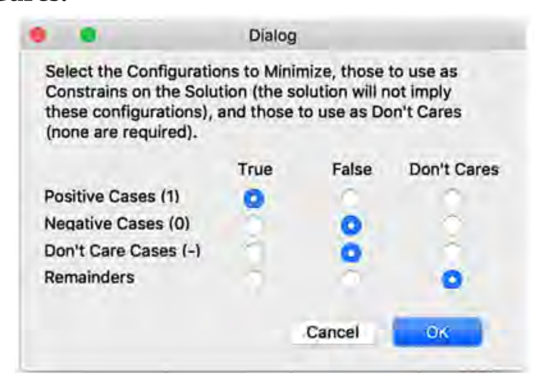
1. Una vez que la *tabla de verdad* ha sido construida, oprima el botón **Specify Analysis** que se ubica en la parte inferior de la ventana:



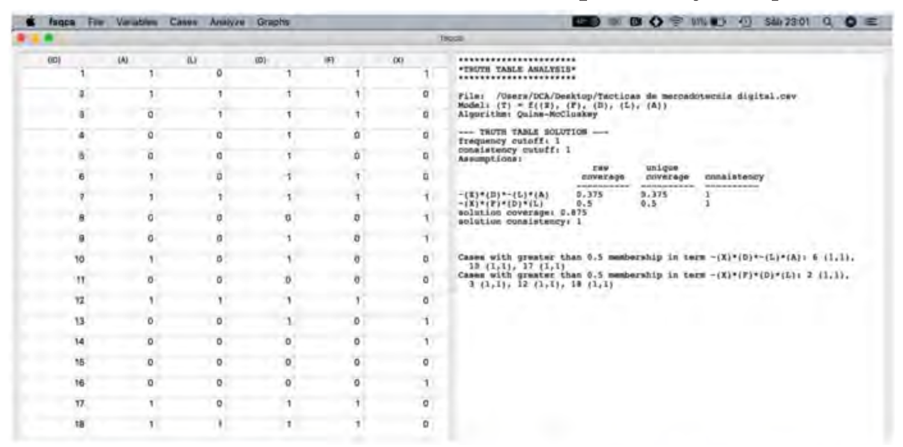
2. En el panel de configuración, marque los casos **Positive Cases** a **True** y todos los demás, en **False**. Esta acción proporcionará *la solución más compleja*. Esta ventana aparece como:

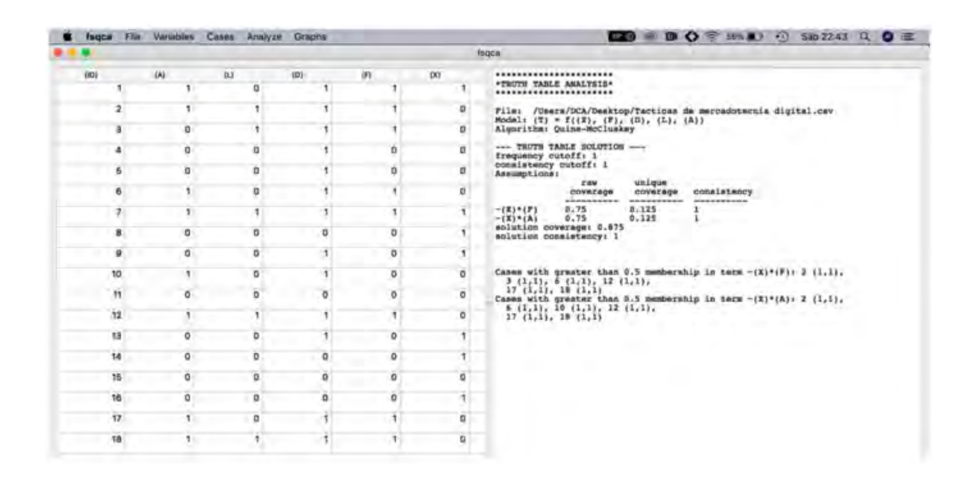


3. Para obtener la solución más *parsimoniosa*, establezca **Positive Cases** en **True**, **Negative Cases** a **False** y **Don´t Cares** así como **Remainders** en **Don’t Cares.**



4. Tenga en cuenta que, cuando el algoritmo para seleccionar implicaciones principales, no puede reducir completamente la *tabla de verdad*, aparecerá la ventana **Prime Implicant Window** y el usuario debe seleccionar los implicaciones principales que se utilizarán, con base al conocimiento del marco teórico de la investigación y el conocimiento empírico sustantivo del investigador. Es muy probable que esta ventana se abra cuando el programa deriva la solución *parsimoniosa*, pero podría suceder para las tres soluciones. (Consulte la sección de análisis de *datos difusos*, para obtener una descripción de cómo funciona esta ventana).
5. Para realizar el análisis, oprima **OK** y la salida aparecerá en la ventana de salida. Se muestran los datos de los casos **puntos 2 y 3** respectivos.

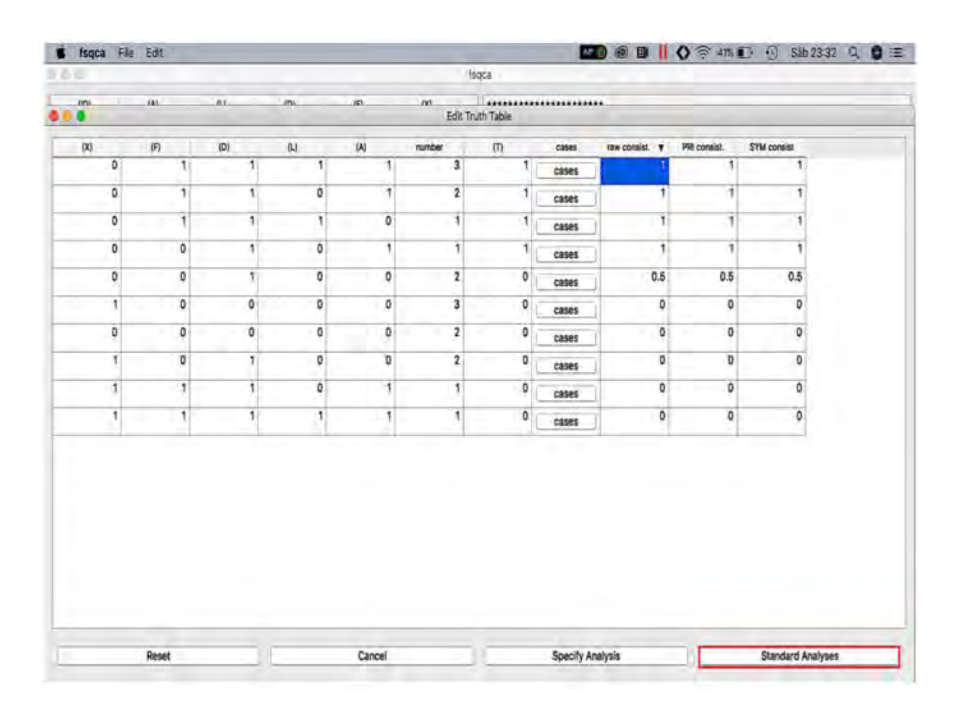




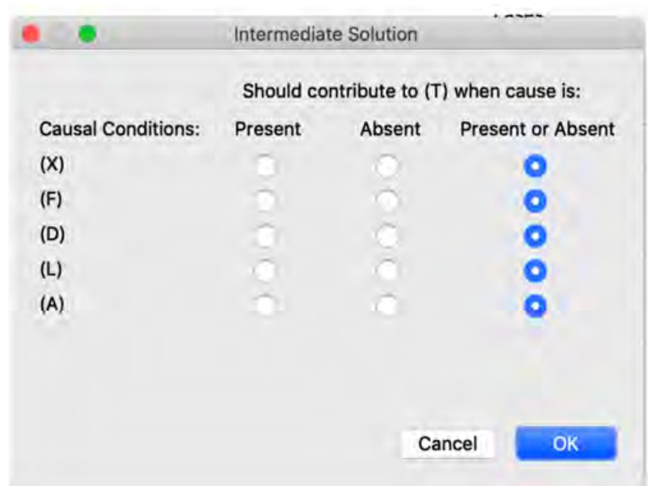
Instrucción: *Standard Analyses Option.*
Caso 5

A continuación se observarán el método más recomendado a utilizar Standard Analyses:

1. Una vez que la *tabla de verdad* esté completamente construida, seleccione **Standard Analyses.** Esta opción, proporciona de forma automática al usuario, soluciones complejas*, parsimoniosas e intermedias*. Este es el procedimiento recomendado a utilizar, ya que es la única forma, por cierto, de obtener la solución intermedia.
2. Para generarla, el software realiza análisis contrafactuales, basados en información sobre las *condiciones causales* que proporciona directamente el investigador.



(X)	(F)	(D)	(L)	(A)	number	(T)	cases	raw consist.	PRI consist.	SYM consist
0	1	1	1	1	3	1	cases	1	1	1
0	1	1	0	1	2	1	cases	1	1	1
0	1	1	1	0	1	1	cases	1	1	1
0	0	1	0	1	1	1	cases	1	1	1
0	0	1	0	0	2	0	cases	0.5	0.5	0.5
1	0	0	0	0	3	0	cases	0	0	0
0	0	0	0	0	2	0	cases	0	0	0
1	0	1	0	0	2	0	cases	0	0	0
1	1	1	0	1	1	0	cases	0	0	0
1	1	1	1	1	1	0	cases	0	0	0



```
**********************
*TRUTH TABLE ANALYSIS*
**********************

File:  /Users/DCA/Desktop/Tacticas de mercadotecnia digital.csv
Model: (T) = f((X), (F), (D), (L), (A))
Algorithm: Quine-McCluskey

--- COMPLEX SOLUTION ---
frequency cutoff: 1
consistency cutoff: 1
                        raw        unique
                     coverage    coverage   consistency
                    ----------  ----------  ----------
~(X)*(D)*~(L)*(A)     0.375       0.375       1
~(X)*(F)*(D)*(L)      0.5         0.5         1
solution coverage: 0.875
solution consistency: 1


Cases with greater than 0.5 membership in term ~(X)*(D)*~(L)*(A): 6 (1,1),
  10 (1,1), 17 (1,1)
Cases with greater than 0.5 membership in term ~(X)*(F)*(D)*(L): 2 (1,1),
  3 (1,1), 12 (1,1), 18 (1,1)
**********************
*TRUTH TABLE ANALYSIS*
**********************

File:  /Users/DCA/Desktop/Tacticas de mercadotecnia digital.csv
Model: (T) = f((X), (F), (D), (L), (A))
Algorithm: Quine-McCluskey

--- PARSIMONIOUS SOLUTION ---
frequency cutoff: 1
consistency cutoff: 1
               raw        unique
            coverage    coverage   consistency
           ----------  ----------  ----------
~(X)*(F)     0.75        0.125       1
~(X)*(A)     0.75        0.125       1
solution coverage: 0.875
solution consistency: 1


Cases with greater than 0.5 membership in term ~(X)*(F): 2 (1,1),
  3 (1,1), 6 (1,1), 12 (1,1),
  17 (1,1), 18 (1,1)
Cases with greater than 0.5 membership in term ~(X)*(A): 2 (1,1),
  6 (1,1), 10 (1,1), 12 (1,1),
  17 (1,1), 18 (1,1)
```

```
**********************
*TRUTH TABLE ANALYSIS*
**********************

File:  /Users/DCA/Desktop/Tacticas de mercadotecnia digital.csv
Model: (T) = f((X), (F), (D), (L), (A))
Algorithm: Quine-McCluskey

--- INTERMEDIATE SOLUTION ---
frequency cutoff: 1
consistency cutoff: 1
Assumptions:
                         raw         unique
                       coverage     coverage     consistency
                      ----------   ----------   ----------
~(X)*(D)*~(L)*(A)      0.375        0.375        1
~(X)*(F)*(D)*(L)       0.5          0.5          1
solution coverage: 0.875
solution consistency: 1

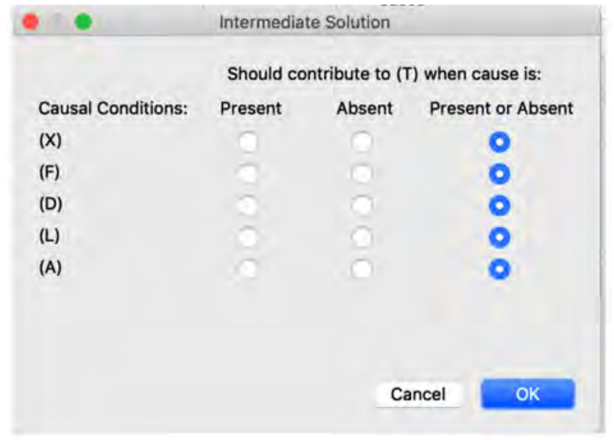

Cases with greater than 0.5 membership in term ~(X)*(D)*~(L)*(A): 6 (1,1),
  10 (1,1), 17 (1,1)
Cases with greater than 0.5 membership in term ~(X)*(F)*(D)*(L): 2 (1,1),
  3 (1,1), 12 (1,1), 18 (1,1)
```

Análisis Standard. Configuración de solución intermedia. Caso 6

Por lo tanto, después de realizar la selección de **Standard Analyses** (**Caso 5**), una ventana que lo guíe a la solución intermedia, aparecerá, por lo que se requiere configurarla.

1. Aquí, el investigador debe seleccionar cómo cada *condición causal* deberá contribuir teóricamente al resultado, como se describió anteriormente, en sus diferentes acepciones:
   a. Si la condición debe contribuir al resultado cuando está presente, seleccione **Present**.
   b. Si la condición debe contribuir al resultado cuando está ausente, seleccione **Absent**.
   c. Si la condición debe contribuir al resultado cuando esté presente **OR** ausente, seleccione **Present or Absent**.

d. Si todas las condiciones están codificadas como **Present or Absent**, entonces la solución intermedia será idéntica a la solución compleja.



2. Tenga en cuenta, que cuando se selecciona el algoritmo de implicaciones principales, no puede reducir completamente la *tabla de verdad*, apareciendo la **Prime Implicant Window** y el usuario debe seleccionar las implicaciones principales que se utilizarán, con base al conocimiento del marco teórico de la investigación y al conocimiento empírico sustantivo. Es muy probable que esta ventana, se abra cuando el programa genere la solución *parsimoniosa*, pero podría suceder para las tres soluciones.
3. Para realizar el análisis, oprima **OK** y las soluciones complejas, intermedias y *parsimoniosas*, aparecerán en la ventana de salida la cual, etiquetará claramente cada una de las soluciones y comenzará por las complejas, luego las *parsimoniosas* y posteriormente presentará, las intermedias.

Resumen para la construcción de tablas de verdad con conjunto de datos nítidos (*crisp-sets*)

Los principios del marco teórico de la investigación, descritos hasta el momento, deben proporcionar la base para todas las técnicas de investigación social, principalmente, las orientadas a la administración de la innovación. Una de las limitaciones de dichos principios teóricos es el que implica solo a conjuntos nítidos de estados Booleanos y, por lo tanto, pueden parecer determinísticos, es decir, absolutos. De hecho, una razón por la cual los científicos de las ciencias sociales orientadas a la administración de la innovación más extremos, desdeñan los análisis basados en el teórico, es la percepción de que está restringido solo a variables de escala nominal.

Una forma de abordar y superar lo anterior, es demostrar que los principios del marco teórico de la investigación, puedan ser aplicados tan bien como conjuntos de *datos difusos*, cuya escala de graduación se hace en valores de rango que van desde **0.0** hasta **1.0**. Como se observa, es una cuestión simple el construir una *tabla de verdad* nítida (*crisp truth table*) de datos dicotómicos. El supuesto aquí, es que el investigador social orientado a la administración de la innovación, tiene simples conjuntos datos binarios, es decir, codificados como **(1)** para presencia y **(0)** para ausencia. Así, los objetivos para la construcción de una *tabla de verdad* nítida (*crisp truth table*), son:

- Examinar la distribución de casos a través de las combinaciones lógicamente posibles de un conjunto dado de *condiciones causales* dicotómicas, y
- Examinar el grado de qué casos con cada combinación de *condiciones causales* están de acuerdo respecto a un resultado dado.

Ver resumen en la **Tabla 3.28.**

Tabla 3.28. Resumen fsQCA para la construcción de tablas de verdad con datos nítidos *(crisp-sets)*

Item	Etapa	Descripción
1	**Creando el conjunto de datos**	Esta etapa se puede lograr usando **fsQCA**, que incluye: **a.** Procedimientos para ingresar datos directamente y para importar conjuntos de datos de otros programas (por ejemplo, archivos delimitados por comas desde **Excel** o archivos delimitados por tabulaciones desde SPSS). **b.** Los archivos importados deben tener nombres de variables simples (sin espacios incrustados o puntuación) como la primera línea del conjunto de datos. **c.** Los datos faltantes deben ingresarse como espacios en blanco.
2	**Ajustes**	Una vez que el conjunto de datos ha sido ingresado o importado y aparece la ventana de la hoja de cálculo de datos: **a.** Haga clic en Analizar (*Analyze*), posteriormente Conjuntos Nítidos (*Crisp-sets*) y finalmente el Algoritmo de la Tabla de Datos (*Truth Table Algorithm*). **b.** Enseguida observará una caja de diálogo, con la etiqueta Seleccionar Variables (*Select Variables*) que estará abierta y que permitirá la especificación del resultado y las *condiciones causales*. **c.** También se puede marcar una casilla para analizar la Negación (*Reverse*) del resultado escogido.
3	**Generación de la *tabla de verdad***	Después de especificar el resultado y las *condiciones causales*: **a.** Haga clic en Ejecute (**Run**), y **fsQCA** generará la *tabla de verdad* completa para el resultado especificado utilizando las *condiciones causales* especificadas. **b.** Una ventana separada se abre en la hoja de cálculo de la *tabla de verdad*.
4	Asignando valores por umbral	Para lograrlo: **a.** El primer elemento de interés es el número de *columna*, que muestra la distribución de casos entre las combinaciones causales. **b.** La *tabla de verdad* es primero presentada con las combinaciones causales ordenadas según la frecuencia, junto con el porcentaje acumulado de casos (que se muestra en el número de *columna*). **c.** La información en esta *columna* debe usarse para seleccionar cualquier umbral de frecuencia que pueda usarse como *valor de corte*. **d.** Cuando el número total de casos en un estudio es pequeño, el umbral debe ser al menos un caso, y las filas de la *tabla de verdad* sin casos (número = 0) debe ser eliminado. **e.** Sin embargo, cuando el número total de casos es grande, se puede usar un umbral más alto para permitir la medición y error de asignación o para generar un análisis de *grano grueso*. **f.** Para borrar filas, simplemente haga clic en la primera fila (la más alta) que se va a eliminar, haga clic en Editar (*Edit*), y posteriormente haga clic en Eliminar fila actual a la última fila (*Delete current row to last row*).

Continuación **Tabla 3.28**

5	**Revisión del puntaje de *consis-tencia***	Para lograrlo: a. El segundo elemento principal de interés es el puntaje de *consistencia* del marco teórico de la investigación, que se muestran en la *columna* de *consistencia* (*Consistency*). b. Con conjuntos nítidos (*crisp-sets*), este cálculo es simplemente la proporción de casos en una fila dada que despliega el resultado en cuestión. c. Un puntaje de 1.0 (o cercano a 1.0) indica alta *consistencia*, es decir, que los casos de la fila coinciden en desplegar el resultado. Un puntaje de 0.0 (o cercano a 0.0) indica que los casos en la fila coinciden en no desplegar el resultado. d. Con conjuntos nítidos (*crisp-sets*), el puntaje de *consistencia* intermedio (0.30 a 0.70) indican que los casos en una fila dada están fuertemente divididos con respecto a la presencia/ ausencia de el resultado. e. Los puntajes de *consistencia* tienen dos usos principales: • Se pueden usar para codificar la *columna* de resultados en la *tabla de verdad*, que se realiza manualmente ingresando 1s y 0s en la *columna* etiquetada con el nombre del resultado, o • Pueden usarse para guiar más investigaciones. f. Ahora bien, suponga por ejemplo, que varias filas tienen puntajes de *consistencia* que indican que los casos son contradictorios, es decir, muchos muestran resultados y muchos no. Identificando estos casos y estudiándolos de cerca, a menudo es posible especificar una *condición causal* que se puede agregar a la *tabla de verdad* a fin de resolver las contradicciones. El investigador puede entonces re-especificar la *tabla de verdad*, incluyendo las condiciones adicionales. Más detalles sobre el refinamiento y el uso de *tablas de verdad* construidas a partir de conjuntos nítidos (*crisp-sets*) pueden ser encontrado en **Ragin (1987)**.

Fuente: Ragin (2008) con adaptación propia.

Glosario

Basado en la obra de **Rihoux y Ragin (2009)**, exponemos el siguiente glosario de términos, esperando le sea de utilidad al lector:

- **Case (*Caso*).** Un caso, es un objeto para el cual, los datos (mediciones) están disponibles, es decir, existen instancias empíricas para mayor referencia (por ejemplo, pozo de petróleo, vehículo terrestre, tipo de vino, paciente, país, pueblo, biblioteca, etcétera). Los casos tratados, son (o deberían ser) más conocidos (o caracterizados), que desconocidos.
- **Causal Combination (*Combinación Causal*).** Es una combinación de las *condiciones causales* o sus complementos (una *configuración*). Si hay ***k*** *condiciones causales*, entonces uno comienza **fsQCA** con $\mathbf{2}^{k}$ combinaciones causales candidatas.
- **Causal Complexity (*Complejidad Causal*).** Mientras más *condiciones causales* sean consideradas, más combinaciones de ellas pueden ocurrir, y por lo tanto, es más compleja la situación.
- **Coherence Data (*Coherencia de Datos*).** Existe coherencia de datos **(Rihoux y Ragin, 2009, p. 15)**, cuando no hay *configuraciones* contradictorias, es decir, no hay casos que sean idénticos con respecto a las *condiciones causales*, pero con diferente en el resultado (*outcome*). Verificar la coherencia de los datos, significa detectar *configuraciones* contradictorias.
- **Condition (*Condición*).** Una condición (también conocida como, variable de condición o *condición causal*, **Rihoux y Ragin, 2009, p. 182**) es una variable explicativa, que puede afectar el resultado (no es una variable independiente variable, en el sentido estadístico).
- **Configuration (*Configuración*).** Una configuración **(Rihoux y Ragin 2009, p. 82)**, es una combinación de *condiciones relevantes* para un resultado (*outcome*) dado; es análogo, a una regla de antecedentes múltiple y puede corresponder a una, más que una, o ningún caso empírico(s).

- **Conjunctural Causation (*Causa por Conjunción*).** La causalidad por conjunción, se refiere a que la causalidad se debe no necesariamente a una sola *condición causal,* pero en cambio, sí a un grupo de *condiciones causales* que se combinan, usando la conjunción con el operador lógico **AND**.
- **Consistency (*Consistencia*).** La *consistencia* **(Rihoux y Ragin, 2009, p. 182)**, es el grado en el cual, la evidencia empírica respalda la afirmación, de que existe una relación teórica establecida basada en el marco teórico de la investigación. Una relación de subconjunto, puede indicar una condición de ser necesaria o ser suficiente; depende sobre cuál es subconjunto: de la causa (*análisis de suficiencia*) o del resultado de interés (*outcome*) (*análisis de necesidad*).
- **Contradictory Configuration (*Configuración Contradictoria*).** La configuración contradictoria **(Rihoux y Ragin, 2009, p. 182)**, es una configuración cuyo resultado (*outcome*) es 1 para algunos casos y 0 para otros casos; por lo tanto, cubre un conjunto de casos empíricos, que, aunque comparten el mismo conjunto de datos de condiciones (por ejemplo, antecedentes), muestran diferentes valores de resultado (*outcomes*) (por ejemplo, consecuentes).
- **Counterfactual Analysis (*Análisis Contrafactual*).** El análisis contrafactual, implica evaluar el resultado (*outcome*) que exhibiría un caso contrafactual si, de hecho, existiera. Es considerar el realizar un *experimento mental* o supuesto. En el análisis contrafactual, los residuos se tratan como combinaciones de no importar (*don't care*); esto da como resultado (a menudo), nuevas combinaciones causales más simples llamadas *contrafactuales*.
- **Counterfactual Case (*Caso Contrafactual*)** (ver también: ***Logical Reminder**, **Difficult Counterfactual Case, Easy Counterfactual Case.*** Un caso contrafactual, es una combinación sustancialmente relevante de *condiciones causales* que no existen empíricamente.
- **Coverage (*Cobertura*).** La *cobertura*, es una evaluación de la forma en que los términos respectivos de las fórmulas mínimas (ver *prime implicants*) cubren los casos observados. Puede haber tres tipos de *cober-*

tura: (1) la *cobertura bruta* (*raw coverage*), que es la proporción de los casos de resultados (*outcome*), que están cubiertos por un término dado; (2) *cobertura única* (*unique coverage*), que es la proporción de casos de resultados (*outcome*), que están cubiertos de manera única por un término dado (ningún otro término cubre esos casos); y (3) *solución por cobertura* (*solution coverage*), que es la proporción de casos que están cubiertos por todos los términos.

- **Difficult Counterfactual Case (*Caso Contrafactual Difícil*).** Los casos de *contrafactuales difíciles*, intentan eliminar una *condición causal* contribuyente de una configuración, mostrando el resultado (*outcome*).
- **Diversity (*Diversidad*).** La diversidad se refiere a si existe o no, un caso para una combinación particular de condiciones de causalidad. En aplicaciones de ciencias sociales, es muy común que no existan casos para muchas combinaciones de *condiciones causales*, y esto se conoce como diversidad limitada.
- **Easy Counterfactual Case (*Caso Contrafactual Fácil*).** Los casos *contrafactuales fáciles*, suponen que, agregar una *condición causal* a una configuración conocida para producir un resultado (*outcome*) (por ejemplo, agregar una condición D a ABC, de modo que el resultado sea ABCD) aún produciría el resultado (*outcome*).
- **Equifinality (*Equifinalidad*).** La equifinalidad, se refiere a diferentes combinaciones causales, que conducen al mismo resultado.
- **Holistic Perspective (*Perspectiva Holística*).** La perspectiva holística **(Rihoux y Ragin, 2009, p. 6)** significa que, cada caso individual es considerado como una combinación compleja de propiedades, un todo específico que no debe perderse o ser minimizado en el curso del análisis.
- **Limited Diversity (ver *Diversidad Limitada*).**
- **Logical Reminder (*Residual Lógico*).** Un *residual* lógico (también llamado caso contrafactual o caso no observado) **(Rihoux y Ragin, 2009, p. 182)** es una configuración (combinación de condiciones) que carece de instancias empíricas.

- **Multiple-Conjunctural Causation (*Causa Múltiple de Conjuntura*).** La frase: *causa múltiple de conjuntura* **(Rihoux y Ragin, 2009, p. 8)**, significa que diferentes trayectorias causales (cada trayectoria causal, es relevante, de manera distinta) pueden conducir, al mismo resultado (*outcome*). El término *múltiple* se refiere al número de caminos, mientras que el término, *de conjuntura,* transmite la noción de que cada trayectoria consiste en una combinación de condiciones.
- **Necessary Condition (*Condición Necesaria*).** Una condición es necesaria para un resultado (*outcome*) **(Rihoux y Ragin, 2009, p. 183)** si siempre está presente, cuando se produce el resultado. En resumen, el resultado es un *subconjunto* de la causa (la misma causa, también puede afectar otros resultados).
- **Net Effects (*Efectos Netos*).** Según **Ragin (2008, pp. 112-114)**, para estimar el efecto neto de un determinada variable independiente, el investigador compensa el impacto de las *condiciones causales* competidoras o contribuyentes, restando, de la estimación del efecto de cada variable, cualquier variación explicada, en la variable dependiente que comparte con otras variables causales. Este es el significado central de los *efectos netos*, el cálculo de la no superposición de la contribución de cada variable independiente, a la variación explicada en la variable dependiente. El grado de superposición, es una función directa de la correlación. Generalmente, cuanto mayor la correlación de una variable independiente sea con sus competidores o contribuyentes, cuanto menor es su efecto neto.
- **Outcome (*Resultado*).** Un resultado (también denominado variable de resultado) **(Rihoux y Ragin, 2009, p. 183)** es la variable a ser explicada por las *condiciones causales*; generalmente, el resultado es el foco principal de un estudio, y es análogo al *consecuente*, en una regla.
- **Prime Implicants (*Implicaciones Principales*).** Las implicaciones principales **(Rihoux y Ragin 2009, p. 183)**, son expresiones reducidas, derivadas en el curso de minimización Booleana. Cada implicación principal, suele ser un conjunto de condiciones unidas por el operador Booleano **AND.** Un subconjunto de las implicaciones principales deri-

vadas, constituye una fórmula mínima, la punto final de una minimización Booleana. Cada implicación principal, en una fórmula mínima, cubre una colección de *configuraciones*, de la *tabla de verdad* para un resultado dado.

- **QCA. Qualitative Comparative Analysis (*Análisis Cualitativo Comparativo*). QCA**, **(Rihoux y Ragin, 2009, pp. xix y xx)** es un término general, que captura tres conjuntos principales (conjunto Booleano, conjunto multivalor y conjunto difuso) como un grupo. Eso común referirse a la versión Booleana original del **QCA** como **csQCA**, donde **cs** denota un *conjunto de datos nítido*, la versión que permite utilizar condiciones de múltiples categorías como **mvQCA**, donde **mv** denota valor múltiple (*multi-value*) y para la versión del conjunto de datos difuso de **QCA** como **fsQCA**, donde **fs** denota el conjunto de datos difuso (*fuzzy-set*).
- **Reminder (*Residual*) (Ver *Logical Reminder*).**
- **Subset Relation (*Relación de Subconjunto*).** En los conjuntos de datos nítidos, una relación de *subconjunto* **(Rihoux y Ragin, 2009, p. 184)** existe entre dos conjuntos, siempre que todos los miembros de un conjunto, estén contenidos dentro del otro conjunto. En los conjuntos de *datos difusos*, se establece que un *subconjunto* existe, en una relación entre dos conjuntos, cuando las puntuaciones de pertenencia en un conjunto, son consistentemente menores o iguales, a los puntajes de pertenencia en el otro conjunto.
- **Sufficient Condition (*Condición Suficiente*).** Una condición (o combinacines de condiciones) es *suficiente* para un resultado (*outcome*) **(Rihoux y Ragin, 2009, p. 184)** si el resultado, siempre ocurre cuando la condición está presente (sin embargo, el resultado puede ocurrir por otras razones también). En resumen, la causa es un subconjunto del resultado.
- **Truth table (*Tabla de Verdad*).** Una *tabla de verdad* **(Rihoux y Ragin, 2009, p. 184)** muestra todas las *configuraciones* (combinaciones de condiciones) unidas por el operador booleano **AND** (intersección).

Referencias

Achen, Ch. (2005). Two Cheers for Charles Ragin… *Studies in Comparative International Development* 40: 27–32.

Bartolini, S. (1993). On Time and Comparative Research. *Journal of Theoretical Politics 5,* (2): 131-167.

Beach, D., y Brun-Pedersen, R. (2013). *Process-Tracing Methods: Foundations and Guidelines*. Michigan: University of Michigan Press.

Berg-Schlosser, D. y De Meur, G. (2009). *Comparative Research Design: Case and Variable Selection,* en Rihoux, B. y Ragin, C. (eds.) *Configurational Comparative Methods.* Thousand Oaks, CA: SAGE.

Berg-Schlosser, D.; De Meur, G., Rihoux, B., y Ragin, C. (2009). Qualitative Comparative Analysis (QCA) as an Approach, en Rihoux,B.; Ragin C. (eds.). Configurational Comparative Methods: Thousand Oaks: CA: SAGE.

Berg-Schlosser, D. (2012). *Mixed Methods in Comparative Politiques: Principles and Applications*. Basingtoke: Palgrave Macmillan.

Bezdeck (1981). Pattern Recognition with Fuzzy Objective Function Algorithms. USA: Kluwer Academic Publishers.

Blackman, T. y Dunstan, K. (2010). Qualitative Comparative Analysis and Health Inequalities: Investigating Reasons for Differential Progress with Narrowing Local Gaps in Mortality. *Journal of Social Policy, 39* (03): 359-371.

Blatter, J. y Haverland M. (2012). *Designing Case Studies: Exploratory Approaches in Small-NResearch*. Research Methods Series. Basingstoke: Palgrave Macmillan.

Bol, D. y Luppi, F. (2013). Confronting Theories Based on Necesary Relations. Making the Best of QCA Possibilities.*Political Research Quarterly 66*(1): 205-210.

Bollen, K. A.; Entwisle,B. y Anderson, A. S.(1993). Macrocomparative Research Methods. *Annual Review of Sociology 19*: 321-351.

Brady, H. E. (4 de abril 2003). *Models of Causal Inference: Going beyond the Neyman-Rubin-Holland Theory*. Artículo presentado en The Annual Meeting of Midwest Political Science Association, Chicago.

Braumoeller, B. (2003). Causal Complexity and the Study of Politics. *Political Analysis* 11: 208–33.

Caren N. y Panofsky, A. (2005). TQCA. A Technique for Adding Temporality to Qualitative Comparative Analysis. *Sociological Methods & Research* 34: 147-172.

Cioffi-Revilla, C. (1981). Fuzzy Sets and Models of International Relations *American Journal of Political Science, 25*, (1): 129-159.

Collier, D. (1998). *Comparative Methods in the 1990s*. Newsletter of the APSA Organized Section in *Comparative Politics 9*, (1): 1-5.

Collier, D. y Mahoney,J. (1996). *Insights and Pitfalls: Keeping Selection Bias in Perspective*. Mimeo: Berkeley.

Coverdill, J. E. y Finlay, W. (1995). Understanding Mills via Mill-Type Methods: an Application of Qualitative Comparative Analysis to a Study of Labor Management in Souhern Textile Manufacturing. *Qualitative Sociology,18* (4): 457-477.

Cronqvist, L. (2005). Introduction to Multi-Value Qualitative Comparative Analysis (MVQCA). *Compasss Didactics Paper*.

Cronqvist, L. (2016) Tosmana versión 1.52, en: https://www.tosmana.net/down-loads/tosmana_manual1_52.pdf

Cronqvist, L. y Berg-Schlosser, D. (2008). *Multi-Value QCA (mvQCA)*, en B. Rihoux y C. C. Ragin (eds.), Configurational Comparative Analysis, Londres: Sage, pp. 69-86.

Curchod, Corentin, Dumez, Hervé, Jeneumaitre, A. (2004). Une étude de de l'organisation du transport aérien en Europe: les vertus de l'aqqc pour l'exploration de la comoplexité. *Revue internationale de politique comparée, 1*(11).

De Meur, G. L.; Rihoux, B. y Yamasaki, S. (2002)*.L'analyse quali-quantitative compare (AQQC-QCA): approache, techniques et applications en sciences humaines.* Louvain-la-Neuve, Academia-Brylant.

De Meur, G. L.; Rihoux, B. y Yamasaki, S. (2009). Addressing the Critiques of QCA, en Rihoux, B. y Ragin C. (eds.). *Configurational Comparative Methods.* Thousand Oaks, CA: SAGE.

Fearon, J. D. (1991). Counterfactuals and Hypothesis Testing in Political Science. *World Politics* 43: 169–95.

Fischer, M. (2011). Social network Analysis and Qualitative Comparative Analysis: Their Mutual Benefit for the Explanation of Policiy Network Structutures. *Methodological Innovations Online, 6* (22): 27-51.

Fiss, P. C. (2010). Building better causal theories: a fuzzy set approach to typologies in organization research. *Academy of Management Journal. 54(2),* en: https://www.researchgate.net/publication/255670704_Building_ Better_Causal_Theories_A_Fuzzy_Set_Approach_to_Typologies_in_Organization_Research/link/56bb757608ae47fa39569f92/download

Gallo, A. (2017). A Refresher on Regression Analysis.*Harvard Business Review ON POINT in The Data-Driven Manager: Maken the Numbers Work for You.* Winter 2017: 22-25.

George, A. (1979). *Case Studies and Theory Development: The Method of Structured, Focussed Comparison*. In *Diplomacy: New Approaches in History, Theory and Policy*, ed. Paul G. Lauren. New York: Free Press.

GII (2019). *Global Innovation Index.* INSEAD-WIPO: Soumitra Dutta, Bruno Lanvin, and Sacha Wunsch-Vincent, en: https://www.wipo.int/edocs/pubdocs/en/wipo_pub_gii_2019.pdf

Gjolberg, M. (2007). The Origin of Corporate Social Responsibility: Global Forces or National Legacies? WPS University of Oslo Centre for Development and the Environment, en: http://www.compasss.org/wpseries/Gjolberg2007.pdf

Goertz, G. (2003). *Assessing the Importance of Necessary or Sufficient Conditions in Fuzzy-sets Social Science* COMPASSS working paper WP2003-7, en: www.COMPASSS.org/wp.htm.

Goldthorpe, J. J. (1997). Current Issues in Comparative Macrosociology: a Debate on Methodological Issues*. Comparative Social Research 16* 1-26

Grofman, B. y Schneider, C. Q. (2009). An Introduction to Crisp-Set QCA with a comparison to Binary Logistic Regresion. *Political Science Quarterly, 62,* (4): 662-672.

Hall, P. A. (2003). *Aligning Ontology and Methodology in Comparative Politics*. En J. Mahoney y D. Rueschemeyer (eds.). *Comparative Historical Analysis in the Social Sciences*. Cambridge: Cambridge University Press 373-404.

Hawthorn, G. (1991). *Plausible Worlds: Possibility and Understanding in History and the Social Sciences*. Nueva York: Cambridge University Press.

Hesse-Biber, S. (2010). *Mixed Methods Research: Merging Theory with Practice*. Nueva York: Guilford Press.

Hicks, A; Misra, J. y Ng, T. N. (1995). The Programmatic Emergence of the Social Security State. *American Sociological Review* 60: 329–49.

Holland, P. W. (1986). Statistics and Causal Inference. *Journal of the American Statistical Association* 81: 945–60.

King, G.; Keohane, R. O. y Verba, S. (1994). *Designing Social Inquiry. Scientific Inference in Qualitative Research*. Princeton, NJ: Princeton University Press.

Klir, G. J.; St. Clair, U. y Yuan, B. (1997). *Fuzzy Set Theory*. Upper Saddle River: Prentice, Hall.

Korjani, M. y Mendel, J. M. (2012). *Fuzzy set Qualitative Comparative Analysis (fsQCA): Challenges and applications.*Conference: Fuzzy Information Processing Society (NAFIPS), en: https://www.researchgate.net/publication/261461558_Fuzzy_set_Qualitative_Comparative_Analysis_fsQCA_Challenges_and_applications

Kraus, S.; Ribeiro-Soriano, D. y Schussler, M. (2017). Fuzzy-set qualitative comparative analysis (fsQCA) in entrepreneurship and innovation research –the rise of a method. *International Management Journal. 14*: 15–33 https://link.springer.com/article/10.1007/s11365-017-0461-8

Lieberson, S. (1991). *Small N's and Big Conclusions: An Examination of the Reasoning in Comparative Studies Based on a Small Number of Cases*, en C. Ragin y H. Becker (ed.), *What is a Case? Exploring the Foundations of Social Inquiry*. Cambridge: Cambridge University Press: 105-118.

Mackie, J. L. (1965). Causes and Conditionals.*American Philosophical Quarterly* 2: 245–65.

Mackie, J. L. (1974). *The Cement of the Universe.* Oxford: Oxford University Press.

Mahoney, J. (2003). *Strategies of Causal Assessment in Comparative Historical Analysis*, en J. Mahoney y D. Rueschemeyer (eds.), *Comparative Historica Analysis in the Social Sciences*.Cambridge: Cambridge University Press, pp. 337-372.

Mahoney, J. y Barrenechea, R. (2017). The logic of counterfactual analysis in case-study explanation. *The British Journal of Sociology:* 1-33.

Mahoney, J. y Goertz, G. (2006). A Tale of Two cultures: Contrasting Quantitative and Qualitative Research. *Political Analysis, 14,* (2): 227-249.

Mahoney, J. y Rueschemeyer, D. (2003) *Comparative Historical Research*. Cambridge: Vambridge University Press.

Marks S. J.; Kumpel E.; Guo J.; Bartram J. y Davis J. (2018). Pathways to sustainability: A fuzzy-set qualitative comparative analysis of rural water supply programs.*Journal of Cleaner Production 205:* 789-798, en: https://www.sciencedirect.com/science/article/pii/S0959652618327410

Marx, A.; Rihoux, B. y Ragin, C. (2014). The Origins, Development and Applications of Qualitative Comparative Analysis (QCA): The first 25 Years. *European Political Science Review, 6*(1): 115-142.

McCluskey, E. J. (1966). Introduction to the Theory of Switching Circuits, Nueva York: McGraw-Hill.

Medina, I.; Castillo, P. J.; Alamos-Concha, P. y Rihoux, B. (2017). *Cuadernos Metodológicos 56*. Madrid: Centro de Investigaciones Sociológicas.

Medzihorsky, J.; Quaranta, M. y Schneider,C. Q. (2020). *Package SetMethods*, en: https://cran.r-project.org/web/packages/SetMethods/SetMethods.pdf

Mejía-Trejo, J.; Sánchez-Gutiérrez, J. y Vázquez-Ávila, G. (2015). Empirical Model for Mobile Learning and their Factors. Case Study: Universities Located in the Urban City of Guadalajara, México. *Revista Apertura, 7* (2):1-13, en: http://www.udgvirtual.udg.mx/apertura/index.php/apertura/article/view/725/498

Mejía-Trejo, J.; Sánchez-Gutiérrez, J. y Maldonado-Guzmán, G. (2016). The Customer Knowledge Management and Innovation. *Contaduría y Administración, 3,* (61): 456-477, en: https://www.sciencedirect.com/science/article/pii/S018610421630002X

Mejía-Trejo, J. (2017a). The Relationship between Competitiveness and Digital Marketing Innovation for a Digital Campaign Design: First Insights Based on a Panel Study in Mexico. *Competition Forum* en: https://www.questia.com/read/1P4-1960315591/the-relationship-between-competitiveness-and-digital

Mejía-Trejo, J.(2017b). The determinant factors of open business model.*Nova Scientia. Revista de Investigación de la Universidad De La Salle Bajío, No. 18*,9 (1): 394-436, en: http://novascientia.delasalle.edu.mx/ojs/index.php/Nova/article/view/710/611

Mejía-Trejo, J. (2017c). Mercadotecnia Digital. Una descripción de las herramientas que apoyan la planeación estratégica de toda innovación de campaña web. México: Editorial Patria, en: https://play.google.com/books/reader?id=AUbJDgAAQBAJ&printsec=frontcover&pg=GBS.PP1

Mejía-Trejo, J. (2019a). *Técnicas Dependientes con SPSS*. México: Alfaomega, en: https://www.alfaomega.com.mx/default/catalogo/profesional/administracion/tecnicas-dependientes-con-spss.html

Mejía-Trejo, J. (2019b). *Diseño de cuestionarios y creación de escalas.Uso del EQS en las ciencias económico-administrativas*. México: Ed. Alfaomega, en: https://www.alfaomega.com.mx/default/catalogo/profesional/administracion/dise-o-de-cuestionarios-y-creacion-de-escalas-uso-del-eqs-en-las-ciencias-economico-administrativas.html

Mejía-Trejo, J.(2019c). How is related the Digital marketing Innovation and e-Leadership in SMEs. Towards a Gender Study. *Colegio de San Luis.* Sep-Dic, 2019: 1-24,en: http://ojs.colsan.edu.mx/ojs/index.php/COLSAN/article/view/948/pdf

Mejía-Trejo, J. (2019d). How the Big Data is Influencing the Open Innovation, First insights into the IT Sector of Mexico. *Acta Universitaria 29:* 1-25, en: http://www.actauniversitaria.ugto.mx/index.php/acta/article/view/1865/ 3016

Mello, P. A. (2014). *Democratic Participation in Armed Conflict. Military Involvement in Kosovo, Afghanistan, and Iraq*. Londres: Palgrave.

Mendel, J. M. y Korjani, M. (2010). Charles Ragin's Fuzzy Set Qualitative Comparative Analysis (fsQCA) Applied to Linguistic Summarization. EUA: Univerisity of Southern California, en: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2517981

Merton, R. K. (1973). *The Sociology of Science: Theoretical and Empirical Investigations.* Chicago: University of Chicago Press.

Mill, J. S. (1967) [1843]. *A System of Logic: Ratiocinative and Inductive.* Toronto: University of Toronto Press.

Morlino, L. (2005). *Introduzione alla ricerca comparata*. Bologna: Il Mulino.

Morse, J. M. (2003). *Principles of Mixed Methods and Multi-Method Research Design*, en Tashakkori, A.; Teddlie, C. (eds). *Handbook of Mixed-Methods in Social and Behavioral Research*. Thousand Oaks, SAGE.

Munck, G. L. (2001). The Regime Question: Theory Building in Democracy Studies.*World Politics, 54*, (1): 119-144.

OECD (2018). Guidelines for Collecting, Reporting and Using Data on Innovation, 4th Ed. París, Francia. Organisation for Economic Cooperation and Developmen (OECD), en: https://www.oecd-ilibrary.org/docserver/9789264304604-en.pdf?expires=1569822203&id=id&accname=guest&-checksum=41982EA3EBE6060AEC51870D0888A774

Pierson, P. (2003). *Big, Slow-Moving, and... Invisible: Macrosocial Processes in the Study of Comparative Politics*, en J. Mahoney y D. Rueschemeyer (eds.), *Comparative Historical Analysis in the Social Sciences*. Cambridge: Cambridge University Press: 177-207.

Przeworski, A. y H. Teune. (1970). *The Logic of Comparative Social Inquiry*. New York: Wiley & Sons.

Quine,W. V. (1952). The problem of simplifying truth functions. *American Mathematical Monthly 59*: 521-531.

Ragin, C. (1987). *The Comparative Method: Moving beyond Qualitative and Quantitative Strategies*. Berkeley: University of California Press.

Ragin, C. (2000). *Fuzzy-sets Social Science*. Chicago: University of Chicago Press.

Ragin, C. (2003a). *Recent Advances in Fuzzy-sets Methods and Their Application to Policy Questions*. COMPASSS working paper WP2003-9, en: www.COMPASSS.org/wp.htm.

Ragin, C. (2003b). *Making Comparative Analysis Count.* Tucson: Mimeo, en: http://www.compasss.org/RaginDayOne.PDF

Ragin, C. (2004). From Fuzzy Sets to Crisp Truth Tables COMPASSS working paper WP2004-28, en: www.COMPASSS.org/wp.htm

Ragin, C. (2006). The Limitations of Net Effects Thinking. In Innovative Comparative Methods for Policy Analysis: Beyond the Quantitative-Qualitative Divide, ed. Benoît Rihoux and Heike Grimm. Nueva York: Springer.

Ragin, C. (2007). *Fuzzy Sets: Calibration Versus Measurement*. Tucson: Mimeo, en: http://www.u.arizona.edu/~cragin/fsQCA/download/Calibration.pdf

Ragin, C. (2008). *Redisigning Social Inquiry. Fuzzy Sets and Beyond.* University of Chicago Press: Chicago y Londres.

Ragin, C. (2009a). *Reflections on casing and case-oriented research*, en D. Byrne and C. Ragin (eds.), *The SAGE Handbook of Case-Based Methods,* Chapter 31: 522-534. Thousand Oaks, CA: SAGE Publications, Inc.

Ragin, C. (2009b). *Qualitative Comparative Analysis Using Fuzzy Sets (fsQCA)* en Rihoux, B. y Ragin, C. (eds.). *Configurational Comparative Methods*. Thousand Oaks. CA: SAGE.

Ragin, C. (2013). New Directions in the Logic of Social Inquiry. *Political Research Quaterly, 66*(1): 171-174.

Ragin, C. y Pennings, P. (2005). Fuzzy Sets and Social Research.Sociological Methods an Research, 33 (4): 423-430.

Ragin, C.; Rihoux, B. (2004). Qualitative Comparative Analysis (QCA): State of the Art and Prospects. *Qualitative Methods,* 2: 3-13.

Ragin, C. y Sonnett. J.(2004). *Between Complexity and Parsimony: Limited Diversity, Counterfactual Cases and Comparative Analysis*. En Vergleichen in der Politikwissenschaft, ed. Sabine Kropp and Michael Minkenberg. Wiesbaden: VS Verlag fur Sozialwissenschaften.

Ragin, C. y Strand, S. I. (2008). *Using Qualitative Comparative Analysis to Study Causal Order*. Comment on Caren and Panofsky (2005), en *Sociological Methods & Research 36*: 431-441.

Ragin, C. y Zaret, D. (1983). Theory and Method in Comparative Research: Two Strategies*, Social Forces 61*(3) 3: pp. 731-754.

Rihoux, B. (2006). Qualitative Comparative Analysis (QCA) and Related Systematic Comparative Methods: Recent Advances and Remaining Challenges for Social Science Research. *International Sociology 21*: 679-706.

Rihoux, B. y De Meur, G. (2009). Crisp-Set Qualitative Comparative Analysis (csQCA) en Rihoux, B. y Ragin, C. (2009). *Configurational Comparative Methods. Qualitative Comparative Analysis (QCA) and Related Techniques*. CA: Thousand Oaks: SAGE.

Rihoux, B. y Lobe, B. (2009). *The Case for Qualitative Comparative Analysis (QCA): Adding Leverage for Thick Cross-Case Comparison*, en Byrne, D. y Ragin, C. (eds). *The SAGE Handbook of case Based Methods.* Londres: SAGE.

Rihoux, B. y Ragin, C. (2009). *Configurational Comparative Methods. Qualitative Comparative Analysis (QCA) and Related Techniques*. CA: Thousand Oaks: SAGE.

Rihoux, B.; Ragin, C.; Yamasaki, S. y Bol, D. (2009). Conclusions.The way(s) ahead, en: Rihoux, B. y Ragin, C. (eds.) *Configurational Comparative Methods. Qualitative Comparative Analysis (QCA) and Related Techniques*. CA: Thousand Oaks: SAGE.

Rihoux B. y Marx, A. (2013). Qualitative comparatibve Analysis at 25 : Stay of the Play and Agenda. *Political Research Quaterly 66*(1): 167-171.

Rihoux, B. y Alamos-Concha, P. (2013). *From Approach to Techniques to Actual Use: What are QCA´s Real Weak Spots?* International QCA Expert Workshop, ETH Zurich. 23-24 octubre 2013.

Rihoux, B.; Alamos-Concha, P.; Bol, D.; Marx, A. y Rezsohazy, I. (2013). From Niche to Mainstream Method?: A Comprehensive Mapping of QCA Applications in Journal Articles from 1984 to 2011. *Political Research Quarterly , 66*(1): 175-183.

Roberts Clark, W.; Gilligan, M. J. y Golder, M. (2006). A Simple Multivariate Test for Assymetric Hypotheses. *Political Analysis, 14*(2): 311-331.

Sartori, G. (1970). Concept Misformation in Comparative Politics. *American Political Science* Review (64): 1033-53.

Seawright, J. (2002). Testing for Necessary and/or Sufficient Causation: Which Cases Are Relevant? *Political Analysis*, *10*(2): 178-193.

Schneider, C. Q. y Grofman, B. (2006). *It Might Look Like a Regression… But It's* Not! An Intuitive Approach to the Presentation of QCA and fs/QCA *Results*. Mimeo. Budapest: Irvine, en: http://www.compasss.org/SchneiderGrofman2006.pdf.

Schneider, C. Q. y Rohlfing, I. (2013). Combining QCA and Process Tracing in Set-Theoretic Multi-Method Research.*Sociological Methods & Research*, *42*(4): 559-597.

Schneider, C. Q. y Wagemann, C. (2003). Improving Inference with a Two Step Approach: Theory and Limited Diversity in fs/QCA.*EUI Working Paper* SPS no. 2003/7, en: https://www.semanticscholar.org/paper/Improving-Inference-with-a-'Two-step'-Approach%3A-and-Schneider-Wagemann/2ced583c51c7ce3228ecc1a670bb56b2541cc905

Schneider, C. Q. y Wagemann, C. (2007a). *Qualitative Comparative Analysis (QCA) und Fuzzy-Sets*. Opladen: Verlag Barbara Budrich.

Schneider, C. Q. y Wagemann, C. (2007b). *Standards guter Praxis in Qualitative Comparative Analysis (QCA) und Fuzzy-Sets*, en S. Pickel, G. Pickel, H.-J. Lauth y D. Jahn (eds.), *Vergleichende politikwissenschaftliche Methoden*. Opladen: WestdeutscherVerlag.

Schneider, C. Q. y Wagemann, C. (2010). Standards of Good Practice in Qualitative Comparative Analysis. *Comparative Sociology 9*: 397-418.

Schneider, C. Q. y Wagemann, C. (2012). *Set-Theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis*. Cambridge: Cambridge University Press.

Skaaning, S-E. (2006). *Democracy besides Elections: an Inquiry into (Dis) respect for Civil Liberty in Latin American and Post-communict Countries after de Third Wave.* Aarhus: University pf Aarhus.

Skocpol, T. (1984). *Emerging Agendas and Recurrent Strategies in Historical Sociology,* en T. Skocpol (ed.), *Visions and Methods in Historical Sociology*. Cambridge: Cambridge University Press: 356-391.

Sobel, M. E. (1995). *Causal Inference in the Social and Behavioral Sciences*. en *Handbook of Statistical Modeling for the Social and Behavioral Sciences*, ed. Gerhard Arminger, Clifford C. Clogg, and Michael E. Sobel. New York: Plenum Press.

Sonnett, J. (2004). Musical Boundaries: Intersections of Form and Content. *Poetics* 32: 247–64.

Stange, K. C.; Crabtree, B. F. y Miller, W. L. (2006). Publishing Multi-method Research. *Annals of Family Medicine, 4*: 292-294.

Tarrow, S. (2010). The Strategy of Paired Comparison: Toward a Theory of Practice. *Comparative Political Studies, 43* (2): 230-259.

Tashakkori, A. y Teddlie, C. (2003). *The Past and Future of Mixed Methods Research: From Data Triangulation to Mixed Model Designs,* en Tasshakori, A. & Teddlie, C. (eds). *Handbook opf mixed Methods in Social and Behavioral Research*. Thousand Oaks, CA: SAGE.

Tetlock, P. E. y Belkin, A. (1996a). *Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives*. Princeton, NJ: Princeton University Press.

Tetlock, P. E. y Belkin, A. (1996b). *Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives*. In *Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives,* ed. Philip E. Tetlock and Aaron Belkin. Princeton, NJ: Princeton University Press.

Thiem, A. y Dusa, A. (2013).Qualitative Comparative Anlaysis with R. A User´s Guide. NY: Springer.

Vanderbrought, Y. & Yamasaki, S. (2003). *The Problem of Contradictory Simplifying Assumptions in Qualitative Comparative Analysis (QCA),* ponencia presentada en la Conferencia general del ECPR, Marburg.

Vaughan, D. (1986). *Uncoupling: Turning Points in Intimate Relationships*. Nueva York: Oxford University Press.

Vink, M. y Vliet, O. (2009). Not Quite Crisp, Not Yet Fuzzy? Assessing the Potentials and Pitfalls of Multi-Value QCA. *Fields Methods*, *21*(3): 265-289.

Vis, B. (2012). The Comparative Advantages of fsQCA and Regression Analysis for Moderately Large-N Analyses. *Sociological methods & Research 41*(1): 168-198.

Wagemann, C. y Memoli, V. (2007). *One or More Approaches to Social Sciences? Different Perspectives on Democracy Support*, ponencia presentada en la Conferencia general del ECPR, Pisa.

Wagemann, C. y Schneider, C. Q. (2007). Standards of Good Practice in Qualitative Comparative Analysis (QCA) and Fuzzy-Sets. *Compasss Working Paper Series.*

Wagemann, C. y Schneider, C. Q. (2010). Qualitative Comparative Analysis (QCA) and Fuzzy-sets: Agenda for a Research Approach and a Data Analysis Technique. *Comparative Sociology* (9): 376-96.

Weber, M. (1949). *Objective Possibility and Adequate Causation in Historical Explanation* en *The Methodology of the Social Sciences*, ed. Edward A. Shils and Henry A. Finch. Glencoe, NY: Free Press.

Web of Science Clarivate (2020). Portal Biblioteca Digital Universidad de Guadalajara, en: https://wdg.biblio.udg.mx/index.php/bases-de-datos

Winship, Ch. y Sobel, M. (2004). *Causal Inference in Sociological Studies*. En Handbook of Data Analysis, ed. Melissa Hardy and Alan Bryman. Londres: SAGE.

Winship, Ch. y Morgan, S. L. (1999). The Estimation of Causal Effects from Observational Data. *Annual Review of Sociology* (25): 659–706.

Yamasaki, S. (2003). *Operationalising and Testing the Concept of 'Policy Domain Profile': Exploiting QCA and Its Intersection Function*, ponencia presentada en la Conferencia General del ECPR, Marburg.

Yamasaki, S.; Rihoux, B. (2009). *A Commented Review of Applications*, en Rihoux, B.; Ragin, C. (eds). Configurational Comparative methods. Thousand Oaks, CA: SAGE.

Zadeh, L. A. (1965). Fuzzy Sets. *Information and Control*, *8*(3): 38–53.

Zadeh, L. A. (1968). Fuzzy Algorithms. *Information and Control. 12*(2): 99-102.

Zadeh, L. A. (1972). A Fuzzy-sets-Theoretic Interpretation of Linguistic Hedges. *Journal of Cybernetics* 2(3): 4–34.

Zadeh, L. A. (1975) The concept of a linguistic variable and its application to approximate reasoning. Information Sciences, Part I: 8, 199-249; Part II: 8, 301-357; Part III: 9, 43-80.



Zapopan, Jal. a 30 de Agosto de 2022

Dictamen de Obra AMIDI.DO.20220830.csQCA

Los miembros del equipo editorial de la Academia Mexicana de Investigación y Docencia en Innovación (**AMIDI**), ver:
https://www.amidibiblioteca.amidi.mx/index.php/AB/about/editorialTeam
se reunieron para atender la invitación a dictaminar el libro:

ANÁLISIS CUALITATIVO COMPARATIVO TOMO I. NÍTIDO (csQCA) .Teoría y Práctica

Cuyo autor de la obra es el **Dr. Juan Mejía Trejo**

Dicho documento fue sometido al proceso de evaluación por pares doble ciego, de acuerdo a la política de la editorial, para su dictaminación de aceptación, ver: https://www.amidibiblioteca.amidi.mx/index.php/AB/procesodeevaluacionporparesenciego

Los miembros del equipo editorial se reúnen con el curador principal del repositorio digital para convocar:

1. Que el comité científico, de forma colegiada, revise los contenidos y proponga a los pares evaluadores que colaboran dentro del comité de redacción, tomando en cuenta su especialidad, pertinencia, argumentos, enfoque de los capítulos al tema central del libro, entre otros.

2. Se invita a los pares evaluadores a participar, formalizando su colaboración.

3. Se envía así, el formato de evaluación para inicio del proceso de evaluación doble ciego a los evaluadores elegidos de la mencionada obra.

4. El comité científico recibe las evaluaciones de los pares evaluadores e informa a el/la (los/las) autor(es/as), los resultados a fin de que se atiendan las observaciones, el requerimiento de reducción de similitudes, y recomendaciones de mejora a la obra.

5. La obra evaluada, consta de:

   introducción, 3 capítulos, glosario y referencias en 305 páginas



6. El desglose de su Contenido, de describe a continuación:

7. Una vez emitidas las observaciones, el requerimiento de reducción de similitudes, y recomendaciones de mejora a la obra por los evaluadores y todas ellas resueltas por el/la (los/las) autor(es/as), el resultado resalta que el contenido del libro:

a. Reúne los elementos teóricos actualizados y prácticos desglosados en cada uno de sus capítulos.
b. Los capítulos contenidos en la obra, muestran claridad en el dominio del tema, congruencia con el título central del libro, y una estructura consistente
c. Se concluye finalmente, que la obra dictaminada, puede fungir como libro de texto principal o de apoyo tanto para estudiantes de licenciatura como de posgrado.

8. Por lo que el resultado del dictamen de aceptación de la obra fue:

FAVORABLE PARA SU PUBLICACIÓN.

Sirva la presente para los fines que a los Interesados convengan.

Atentamente

Dr. Carlos Omar Aguilar Navarro.
ORCID: https://orcid.org/0000-0001-9881-0236
Curador Principal AMIDI.Biblioteca
AMIDI